

SEARCH REQUEST FORM
Scientific and Technical Information Center

Access DB#

53593

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: _____	NA Sequence (#) _____	STN _____
Searcher Phone #: _____	AA Sequence (#) _____	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: _____	Bibliographic _____	Dr. Link _____
Date Completed: <u>10/30/01</u>	Litigation _____	Lexis Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems _____
Client Prep Time: _____	Patent Family _____	WWW Internet _____
Online Time: _____	Other _____	Other (specify): _____

***Contact:**
Sheppard

PT 308-4499

tel: 308-4499

~~STIC~~-Biotech/ChemLib

53593

From: Prouty, Rebecca
Sent: Wednesday, October 24, 2001 8:44 AM
To: STIC-Biotech/ChemLib
Subject: Sequence Search

Art Unit 1652, 10A13
Mailbox: 10C01
308-4000
Serial Number: 09/416,579

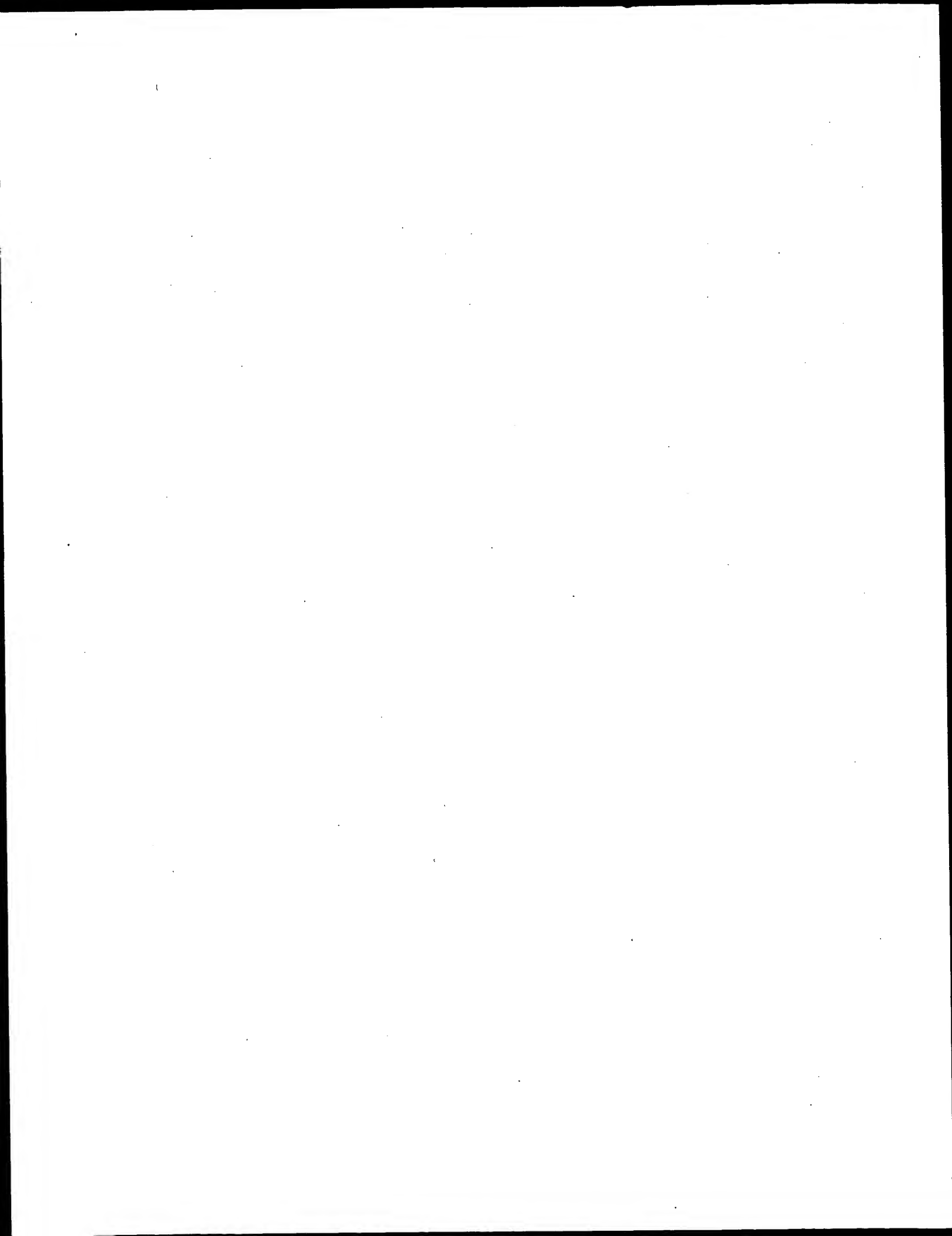
Please search and interference search SEQ ID NO:1 and the translation of SEQ NO:1

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, **.rnpm** and **.rnpn**

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, **.rapm** and **.rapn**

The Pending database search results should not be left in the case because they contain data that is confidential.



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 03:57:24 ; Search time 1156.6 Seconds
(without alignments)
6154.244 Million cell updates/sec

Title: US-09-416-579A-1

Perfect score: 753

Sequence: 1 atggcggagggagcagctctg.....gccagagggtccgacagataa 753

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

```

```

190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Downloaded from <http://ajphaphysocpharm.sagepub.com/> at 11:06 11 November 2014

JOURNAL	Unpublished (1997)
COMMENT	Contact: Harvey, D.
	G. M. Rubin-Molecular and Cell Biology
	University of California Berkeley
	539 USA, Berkeley, CA 94720-3200, USA
	Fax: 510 643 9947
	Email: http://www.fruitfly.org/BST , est@fruitfly.berkeley.edu
	Plate: 413 row: C column: 6
	High quality sequence stop: 458.
FEATURES	
source	Location/Qualifiers
	1..565
	/organism="Drosophila melanogaster"
	/db_xref="taxon:7227"
	/clone="LD41330"
	/clone_lib="ID Drosophila melanogaster embryo pot2"
	/sex="male and female"
	/dev_stage="0 to 24 hours mixed stage embryonic"
	/lab_host="XLI Blue"
	/note="Organ: embryo; Vector: pOT2; Site_1: EcoRI; Site_2: XhoI; Sized fractionated cDNAs were directly ligated into pOT2."
BASE COUNT	149 a 159 c 146 g 111 t
ORIGIN	
Query Match	61.8%; Score 465.2; DB 21; Length 565;
Best Local Similarity	99.4%; Pred. No. 1.5e-115;
Matches 467; Conservative	0; Mismatches 3; Indels 0; Gaps 0.
Oy	1 atgagcgagacatcctgtgccgaagaaggagacaagtacgcgcgaaggaccagccc 60
Dd	96 ATGGCGAGGACACATCTGTGCCGAAGAAGGGAGACAAGTACGCCGAGGACCAACC 155
Oy	61 ttcaagtcctcatcgaaaggacaatcggcagcgggaagaaaccaagttattgaacctt 120
Dd	156 TTCACTCCTTCCTCATCCAGGGCAACATCCGCACCGGGAAGACAGTATTGGAACCACTTC 215
Oy	121 gagagatcacagaagacacttgcctgcgtgcgaccgagccgcgcgaagaattgacgaagtc 180
Dd	216 GAGAAGTACAGAACGACATTGCTGCTGACCCGAGCCCTTGAGAAATGGCGCAACGTC 275
Oy	181 aaacgggtaaatctgtcgtgagctgatgttcaacaaga tcccaagaag tgggcatgcctt 240
Dd	276 AACGGGTAATATCGTGGAGGTGATGATCAAAAGATCCCAAGAAAGTGGCCATGCCCTTT 325
Oy	241 cagagttatgtcacgctgacatgctgcagttgcgacacccgcccaacaagaagcta 300
Dd	336 CAGAGTTATGTCAACGGTGACATGATGCGACAGTCGACACCGTCCCAACCAACAAACTTA 365
Oy	301 aaataargagagcttccacttttagcgtgcgtacttgcttcgtggagaaatatgcagca 360
Dd	366 AATAARGAGAGCTTCCACTTTTAGCGTGCCTTGCCTTCGTTGGAGAAATATGCTCCCTA 425

[illegible]

	Df	509	AACGCGCTCCCTGGAGCAGGAGCATATGCATTACGTACAGTTCAATCGAA	568
	OY	421	gagtcattcactgtgaaggcggaacctcatatatactctgcacactgcggaaagtgcg	480
	Df	569	GAGTCATTACAGTGCAGCGGACCTCATATATCTCGCACCTCGCGAAGTGCGG	628
RESULT		5		
LOCUS		AA441228	590 bp	mRNA EST 27-NOV-1998
DEFINITION		LD15983.5prime LD Drosophila melanogaster embryo Bluescript		
ACCESSION		AA441228	Drosophila melanogaster cDNA clone LD15983 5prime, mRNA sequence.	
VERSION		AA441228.1	GI:2794077	
KEYWORDS		EST.		
SOURCE		fruit fly.		
ORGANISM		Drosophila melanogaster		
REFERENCE		Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
AUTHORS		Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;		
		Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
		1 (bases 1 to 590)		
		Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C., Broksstein		
		,P., Lewis,S. and Rubin,G.M.		
TITLE		BDDP/HMI Drosophila EST Project		
JOURNAL		Unpublished (1997)		
COMMENT		On Jan 19, 1998 this sequence version replaced gi:2153106.		
		Contact: Harvey, D.		
		G. M. Rubin-Molecular and Cell Biology		
		University of California Berkeley		
		539 LSA, Berkeley, CA 94720-3200, USA		
		Fax: 510 643 9947		
		Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu		
		Plate: 159 row: G column: 11		
		High quality sequence stop: 412.		
FEATURES		Location/Qualifiers		
source		1..590		
		/organism="Drosophila melanogaster"		
		/db_xref="BDDP_EST:BDCl015118"		
		/db_xref="taxon:7227"		
		/clone="LD15983"		
		/clone_lib="LD Drosophila melanogaster embryo Bluescript"		
		/sex="Male and female"		
		/dev_stage="0 to 24 hours mixed stage embryonic"		
		/lab_host="SOLR"		
		/note="Organ: embryo; Vector: Bluescript SK; Site:1; EcoRI		
		/site:2: XhoI; Constructed using Stratagene Zap-CNA		
		Synthesis kit. Oligo dT-primed and directionally cloned at		
		EcoRI and XhoI in Bluescript SK(+/-)"		
BASE COUNT		158 a 165 c 150 g 117 t		
ORIGIN				
Query Match		59.3%	Score 446.2;	DB 7; Length 590;
Best Local Similarity		99.1%;	Pred. No. 2.1e-110;	
Matches 459; Conservative		0; Mismatches 3;	Indels 1; Gaps 1;	
OY	1	atgagggagagcaatccctgtcgcgaaaaggagaccagtaagccgagagaccagccc	60	
Df	128	ATGGCGGAGGAGCATCTGTGTGCCGAAGAGGTACCAAGTACCGGAGCACCAGCCC	187	
OY	61	ttaacgctccatctagaaggacaacatcgcgacgaggaaaccagctatttgaaaccttc	120	
Df	188	TTTACCCGTCATTCATGAGGCGAACATCGCGCAGCGGAGGAGACACGATTTGAACCATTC	247	
OY	121	gagaagtaacaagaacatttgcctgtctgttagccgagccgfcggaagtgtgcgaagtc	180	
Df	248	GAGAAGTACAAGAACGACATTTCCTGCTTGACCGAGCCCGTGAAGAGTGCGCAACGTC	307	
OY	181	aacggggtanaactctgcttgaagctgatgtacaaaagatccccaaagaattggccatgccctt	240	
Df	308	AACGGGTAATCTGCTGTGAGACTGATGTACAAAGATCCCAAGAGTGGCCATCCCTTT	367	

Query Match	19.2%	Score 144.4	DB 30	Length 689
Best Local Similarity	57.6%	Pred No. 1.3e-28		
Matches 296	Conservative	0	Mismatches 213	Indels 5
				Gaps 2
QY 158	ccgcgcgaagatggcgcgaacgctcaacggggtaaatctgtcgtgagcgtatgtacaagaatc	217		
Db 1	ccgttgaaatgtgcgcagatctttaaagggttgcaatctttttggaacatcatgacaaagatc	60		
QY 218	ccaaagaatgggcacatgc--ccttcaaggtttgtcgaagctacacatctgcagtcgca	275		
Db 61	ccgaaaaaattggggcgatgaacattccagtcantacgtttcccttgacgattgttgacatgca	120		
QY 276	caccgcccacccaacaagaagctaaataaayagagcgtctccatctttagcgtcgcgta	335		
Db 121	ccggagaccctgcgtccaaactccactgaaagctaatgagacgcattatattcagtcgagata	180		
QY 336	ttgcttcgttgaggaacatctgcgaagaacggcctgcgtcggagcgaagcattgacaatacgtc	395		
Db 181	ctctcttcgtttggaacacattgattgaaataataaacattccatccagcagcttgcagattct	240		
QY 396	ggaggagatgttaacagttcaatcgaagaagctcatctacactcgaagcggaccatcatata	455		
Db 241	tgatgatgtgttccnnttcatccacacacacattccattgatgactgatttgatagata	300		
QY 456	tctgcgaccctgcgcggaggctggcgtacgaacgcatctcggcagcgggctgcttcaggga	515		
Db 301	tcttaaacatcatcacttcaattagatggtaccaaagatcaaaagagacctggttcagaga	360		
QY 516	gaagctggttcgcgttaagaaccttcgaagagctgacatgattttgaccaggagactgttgat	575		
Db 361	gcagtgtgtccctctgcatatcatttgaggaactgcataagttgcatagagactggcattat	420		
QY 576	acacccagagacgaaccggaat---cggtcaagttcctagttccatgcgatctgaaact	632		
Db 421	caacagagattacatcctgaattgtcccgacagagatggttgaattgatact	480		
QY 633	ggaaaacatctggcaccgagatcacagcgtcctggag	666		
Db 481	ctctcagatgaaccgatgaataacaaagaaagtga	514		
RESULT 9				
LOCUS	AL583655	986 bp	mrna	EST 16-FEB-2001
DEFINITION	AL583655 L7T_NFL006_P12	Homo sapiens	cdna clone	CS0DK008YA23 5
ACCESSION	AL583655			
VERSION	AL583655.1	GI:12952825		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
TITLE	1 (bases 1 to 986)			
JOURNAL	Li, W. B., Gruber, C., Jesse, J. and Polayes, D.			
COMMENT	Full-length cDNA libraries and normalization			
	unpublished (2001)			
	Contact: Genoscope			
	Genoscope - Centre National de Sequencage			
	BP 191 91006 Evry cedex - France			
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.			
FEATURES	Location/Qualifiers			
source	1..986			
	/organism="Homo sapiens"			
	/db_xref="taxon.9606"			
	/clone="CS0DK008YA23"			
	/clone_id="L7T_NFL006_P12"			
	/tissue_type="placenta"			
	/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA			
	was primed with a NotI-0.10kg(dt) primer. Five prime end			
	enriched, double-stranded cDNA was digested with Not I and			

Db 510 GATCATTTTACAGGCCAAGATACATATTTTGGGAAATCTTTATMAAAGTGGAAAGATGC 565

QY 374 agcaggagcgtlacatacgccttgagagatggtlcaacgtlcatcgaagtcgacat 428

Db 570 CCGAGGTGGACTTCGCTGCTTTAAGTGAATGGTTTGCATGGAATATCAAGAACAT 624

RESULT	13
BF652168	
LOCUS	BF652168 294 bp mRNA EST 20-DEC-2000
DEFINITION	Z75654 MARC 3BOV Bos taurus CDNA 5' , mRNA sequence.
ACCESSION	BF652168
VERSION	BF652168.1 GI:11917298
KEYWORDS	EST .
SOURCE	COW .
ORGANISM	Bos taurus

Db 178 ACCATGCTCGACCGACACACTCGTCGCCAGACGTTGCCCTGTACCGTTGATGGAGCGGTCA 237

Qy 319 attctttagcgtcgcgtatgtcttcgttgagacatcgacgaacgcg 365

Db 238 ATTTCAGTGCACAAGTACGTTTGTGTAAGAAACTTGTACAGAAATGG 264

RESULT	14
AM226818	
LOCUS	531 bp mRNA
DEFINITION	un662f04.y1 Sugano mouse kidney mRna Mus musculus cDNA clone IMAGE:2285583 5' similar to TR:O15238 O15238 THYMIDINE KINASE ; mRNA sequence.
ACCESSION	AM226818
VERSION	AM226818.1 GI:6556114
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGON R;
 RX MEDLINE=20158988; PubMed=10692477;
 RA Munch-Petersen B., Knecht W., Ienz C., Sondergaard L., Piskur J.;
 RT "Functional Expression of a Multisubstrate Deoxyribose Kinase
 from *Drosophila melanogaster* and Its C-terminal Deletion Mutants";
 RL J. Biol. Chem. 275:6673-6679(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Munch-Petersen B., Sondergaard L., Piskur J.;
 RT Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AE003724; AAF55615.1; -
 DR EMBL: AF045610; AAD47355.2; -
 DR EMBL: AF185268; AAD56545.1; -
 DR EMBL: Y18048; CAB41881.1; -
 DR FlyBase: FBgn0022338; dmk.
 DR InterPro: IPR002624; -
 DR Pfam: PF01712; dmk; 1.
 KW Kinase.
 SQ SEQUENCE 250 AA; 29088 MW; 17C5DF197B8792DB CRC64;

alignment_scores: Quality: 1290.00 Length: 250
 Ratio: 5.202 Gaps: 0
 Percent Similarity: 99.200 Percent Identity: 98.400

Alignment block:

US-09-416-579A-1 x Q9XZT6 ..

Align seg 1/1 to: Q9XZT6 from: 1 to: 250

1 ATGGCGAGGAGCAGCATCTGTGCCGAAAGGAGGAGCAAGTACGCCGAGG 50
 1 MetLaGlunLaIaIaSerCysAlaArgLysGlyThrLysTyrLaGlunG 17
 51 CACCCACCCCTTCACCGCTCTCATCGAGGGCAACATCGGAGCGGAGAGA 100
 17 ThrGlnProPhetThrValLeuIleGlnGlyAsnIleGlySerGlyLys 34
 101 CCAAGTATTGAACCACTTCGAGAGTACAGAGCAAGCAATTCCTCCGCTG 150
 34 hrThrTyrLeuAsnHisPheGlnLysTyrLysAsnSplIleCysLeuLeu 50
 151 ACCGAGCCCGTCGAGAGTGGCCCAACATTCACGGGGTAATTCGTGGA 200
 51 ThrGlnProValGlnLysTyrPArgAsnValAsnGlyValAsnLeuLeuG 67
 201 GCGATGTACAAAGATCCCAAGAGTGGCCATGCCCTTCAGAGTTATG 250
 67 uLeuMetTyrLysAspProLysTyrPArgMetProPheGlnSerTyr 84
 251 TCAGCGTGCATGTCGAGTGCACACCGCCCAACCAAGAGGTA 300
 84 alThrLeuThrMetLeuGlnSerHisThrAlaProThrAsnLysLysLeu 100
 301 AAAATTAAGAGCGCTCATTTTAAAGCTGCTATTCCTGCGAGAA 350
 101 LysIleMetGlnArgSerIlePheSerAlaArgTyrCysPheValGlnLys 117
 351 CATGCGAGAAAGCGCTGCGAGAGGCGGCAATGATACCTGCGAGAG 400
 117 mMetArgAlaGlnGlySerLeuGlnGlnGlyMetTyrAsnThrLeuGlnG 134
 401 AGTGTTACAGTTTCATGGAAGATCCATTCACCTGAGGCGAGCTCATC 450
 134 lntTyrTyrLysPheIleGlnGlnSerIleHisValGlnAlaLysPheLeu 150
 451 ATATATCTGCGACCTCGCCGAGAGTGGCGTACGAACGATCCGCGAGCG 500
 151 lIeTyrLeuArgThrSerProGlnValAlaTyrGlnArgIleArgGlnArg 167

501 GCGTCGTTCTGAGGAGAGCTGCGCCGCTTAAAGTACCTTCAGAGACTGC 550
 167 gAlaArgSerGlnGlnLysSerCysValProLeuLysTyrLeuGlnGlnLeu 184
 551 ATAGTTGACACGACGTGTTGATACACGAGAGACGAGCGAGCTGTC 600
 184 lStLeuHisGlnAspTyrPheIleHisGlnArgArgProGlnSerCys 200
 601 AAGTCTTAGTCTTCGATGCCGATTCGAACCTGGAACATTCGGACCGA 650
 201 LysValLeuValLeuAspAlaAspLeuAsnLeuGlnAsnIleGlyThrG 217
 651 GTACGAGCGCTCGAGAGACGATATTCACGCCCATCTTAAGTAAACAC 700
 217 uTyrGlnArgSerGlnSerIlePheAspAlaIleSerSerAsnGlnG 234
 701 AGCCCTCGCCGTTGCTGTCGCCGACGACGACGAGGCTGCCCGA 750
 234 lntProSerProValLeuValSerProSerLysArgGlnArgValAlaArg 250

seq_name: sp_virus:Q9J579

seq_documentation_block:

ID Q9J579 PRELIMINARY; PRT; 235 AA.
 AC Q9J579;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 DE ORF FPV151 DEOXYCYTIDINE KINASE.
 GN FPV151.
 OS Fowlpox virus.
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Avipoxvirus.
 OX NCBL_TaxID=10261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
 RX MEDLINE=20193820; PubMed=10729156;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF198100; AAF44495.1; -
 DR InterPro: IPR002624; -
 DR Pfam: PF01712; dmk; 1.
 KW Kinase.
 SQ SEQUENCE 235 AA; 27678 MW; 60AC94B9B0B956D4 CRC64;

alignment_scores: Quality: 335.00 Length: 212
 Ratio: 2.359 Gaps: 5
 Percent Similarity: 66.981 Percent Identity: 36.792

Alignment block:

US-09-416-579A-1 x Q9J579 ..

Align seg 1/1 to: Q9J579 from: 1 to: 235

67 CTCCTCATCGAGGCGACATCGGAGCGGAGAGCAAGCATTTGAACCA 116
 26 lIleSerIleGlnGlnLysAsnIleSerAlaGlyLysSerThrLeuIleAsn 42
 117 CTTCGAGAGTACAGAGACATTTGCTGTCGACCGAGCGGTCGAGA 166
 42 eLeuSer.....AspAsnGlyTyrAsnValValGlnGlnProLeuGlnG 57
 167 AGTGCGCAACGTCACGCGGTAATCTGAGAGCGAGTGAACAGT 216
 57 lntPArg.....GlyAsnAsnLeuLeuAspLysLeuTyrLysAsp 70

```

217 CCCAAGAGTGGCCATGCCCCCTTCAGAGTTATGTCACCGCTGCATGCT 266
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
71 ProseriaterpalaTythrphglnSerHisAlaThrTrpThrargTh 87
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
267 GCAGTCCACACGCGCCCAACCAAGAGCTAAAG.....ATTAARG 310
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
87 rLySTHrTyrIleAspAlaLeuAsnLysLysGlyAsnIleIleLeuG 104
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
311 AGCGCTCCATTTTAAAGCGCTGCTATGCTTCGTGAGAACATGACGCA 360
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
104 LurArgSerValrPheSerAspLysTyrIlePheAlaThrAlaLeuHisAsp 120
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 AACGCTCGCTGGACGACGACATGTACATACCGCTGGACAGCTGTACAA 410
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 IleLysTyrIleAspAspThrIleGlyTrpAsnIleTyrAsnGlyTyrSer 137
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
411 GTTCATGAGAGTTCATTCACCTGACGCGGACCTCATCATATATCTGC 460
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 STPrMetThrGluPheMetAspIleLysIleAspGlyIleIleTyrLeuL 154
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
461 GCACCTCGCGGAGAGTGGCGTACGACGACATCGGCGACGGGCTGTCT 510
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
154 ySTrSerProAspIleCysTyrLysArgMetLeuAsnArgAlaArgHis 170
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
511 GAGAGAGCTGCGCGCGCTTAAGTACCTTCAGGAGCTGCATGAGTTGCA 560
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
171 GluLysAsnThrValLysIleAspTyrLeuAsnLeuLeuHisAspLysH 187
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
561 CCAGAGACTGTTGTTATACACGACGACGCGCATGCTGCAAGGCTCTAG 610
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 sGLyLysTrpLeu...SerGluAsnAsnGlnLysGluPheLysValrPro 203
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
611 TCCTCGATGCGGATCGAATCGAATAACATTTGGCACCAGAGTACCAGCC 660
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
203 aLLeuGluIleAsnGlyAspGlyAspPheIleAspAspSerAsnArgGln 219
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
661 TCGGAG.....AGCAGCATATTCAGCGCCATCTCA 690
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 SerSerIleLeuSerAsnIleTyrAspPheIleSer 231

```

seq_name: sp_virus:055749

```

seq_documentation_block:
ID 055749 PRELIMINARY; PRT; 195 AA.
AC 055749;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DE 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE PUTATIVE THYMIDINE KINASE.
OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=10488;
RN [1]
RP SEQUENCE FROM N.A.
RA Bahr U., Tidona C.A., Darai G.;
RL Virus Genes 0:0-0(1997).
DR EMBL: AF003534; AAB94460.1; -.
DR InterPro: IPR002624; -.
DR Pfam: PF01712; dmk.1.
SQ SEQUENCE 195 AA; 22822 MW; 7F0228057F2CB85D CRC64;

```

alignment_scores:

Quality:	187.00	Length:	195
Ratio:	1.545	Gaps:	7
Percent Similarity:	62.051	Percent Identity:	27.652

alignment_block:

us-09-416-579a-1 x 055749 ..

Align seg 1/1 to: 055749 from: 1 to: 195

```

67 GTCCTATCGAGGCGCAACATCGGACGCGGAAGACACGATATTTGAACA 116
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4 IleCysIleAspGlyIleIleGlyAlaGlyLysSerThrValrThrHisAr 20
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
117 CTTGAGAGAG.....TACAGAGACGACATTTGCTGTCAGCCGACCCG 160
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 gleuLysLysAsnLeuTyrLys.....CysTyrGluGluProI 33
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
161 TCGAGAGAGTGGCGCAACGTCACAGGGGTAATCTGCTGAGAGCTGATGAC 210
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
33 LeAspLysTrp.....ThrLeuLeuProAsnLeuTyr 43
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
211 AAAGATCCCAAGAGAGTGGCGCATGCGCTTTCAGAGTTATGTCACCGCTGAC 260
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
44 AsnAspMetLysLysTyrAlaThrProPheGlnProPheGlnValLeuPhe 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
261 CATGCTGCAGTGCACACCGCCCAACAC...AAGAGCTAAAAATTA 307
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 rGlnTyrAspGlnTyrLeuSerPheLysAspCysLysGlyTrpValrValr 77
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
308 KGGAGCGCTCCATTTTAAAGCGCTATGCTTCGTGAGAACATGCGCA 357
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
77 aLLeuArgCysProIleTrpThrSerLysAsnIlePheThrSerLeuMetIle 93
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
358 CGAAGCGCTCGCTGCGACGCGCATGTACAAATACGCTGAGAGAGTGTGTA 407
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
94 GluAsnAsnLeuPheAspLeuSerAlaIleAspThr.....Tyr 106
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
408 CAAGTTCATCGAAGAGTCCATTCACCTGACGCGGACCCATCATATATAC 457
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 rAsnAsnLeuTyrGluArgLeuSerTyrGlnValrAspHisPheIleTyr 123
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
458 TGGCAGCTTCGCGGAGTGGCGTACGACGACATCGGCGGCGGCTGCT 507
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
123 IleLysValrAspSerGluMetAlaPheGluArgIleLysLysArgAspArg 139
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
508 TCTGAGAGAGCTGCTGCGCGCTTAAGTACCTTCAGAGCTCATGAGTT 557
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
140 PheAlaGluGlnAsnIleSerPheAspTyrLeuLysSerLeuGlnLys 156
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
558 GCACGAGAGCTGTTGATACACGACGACGACGCGGCTGTCAGAGTCC 607
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
156 sTyrAlaThrSerLeuAla...ThrLeuSerProSerThr.....Valr 170
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
608 TAGTCTCGATGCGGATCGAATCGAATAACATTT 642
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
170 hrIleIleAspGlySerAsnThrIleGluGluVal 181

```

seq_name: sp_virus:098468

```

seq_documentation_block:
ID 098468 PRELIMINARY; PRT; 188 AA.
AC 098468;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
DE 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE SIMILAR TO BACILLUS SUBTILIS 24.1 KDA PROTEIN.
GN A416R.
OS Paramesitium bursaria chlorella virus 1 (PBCV-1).
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phycodnavirus.
OX NCBI_TaxID=10506;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96400190; PubMed=8806566;
RX Kutish G.F., Li Y., Lu Z., Furuta M., Rock D.L., Van Etten J.L.;
RT "Analysis of 76 kb of the chlorella virus PBCV-1 330-kb genome: map
  positions 182 to 258."
  Virol J 223:303-317(1996).
RL
DR EMBL: U42580; AAC96784.1; -.
DR InterPro: IPR002624; -.
DR Pfam: PF01712; dmk.1.
SQ SEQUENCE 188 AA; 22361 MW; 82CCTA4AF3F55F2A CRC64;

```



```

672 CATATTCGACGACCATCTCAAGT.....AACCAAC 7000
|||||
200 u1ethrlysl1lemtsetseraProlySglyLysAsnPhelIeasnProG 217
|||||
701 AGCCCTCG 708
|||||
217 InProser 219

seq_name: sp_virus_Q9YTP5

seq_documentation_block:
ID Q9YTP5 PRELIMINARY; FRT; 527 AA.
AC Q9YTP5;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE THYMIDINE KINASE.
OS Atelein herpesvirus 3.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=85618;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=73;
RA Albrecht J.-C., Fleckenstein B.;
RT "Primary Structure of the Herpesvirus Ateles Genome.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF083424; AAC95545.1; -.
DR InterPro: IPR000531; -.
DR Pfam: IPR001889; -.
DR Pfam: PF00693; TK_herpes; 1.
DR PRODOM: PD001519; -.
DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW kinase.
SQ
SEQUENCE 527 AA; 59674 MW; 68CE835449FOA0BC CRC64;

alignment_scores:
Quality: 149.50 Length: 254
Ratio: 1.133 Gaps: 10
Percent Similarity: 51.969 Percent Identity: 25.197

alignment_block:
US-09-416-579A-1 x Q9YTP5 ..

Align seg 1/1 to: Q9YTP5 from: 1 to: 527

40 TAGCCGAGGGGACACCCAGCCCTTCAGCGTCCATCGAGGACCAATCGG 89
::: ||| ||| ::::: |||||::|||
203 HIsGInGluProIleProAlaPhaPheH1ephLeuGluGlySerIleG1 219
90 CAGCGGGAGAACCGATTTGAACCACTTCGAAATAC.....A 130
| ||||| ||||| ||::: :::::
219 yAlaSlYslYsthrTrhLeuLeuLysSerMetAlrGlyIleLeuProGlyL 236
131 AGAAGACATTTGGCTGCTGACACGAGCCGCGAGAGATGGCGCAACGTC 180
||||| :::: ||| ::::: |||||::|||
236 yAsn...ValLeuThrPheH1stGluProMetAlaPheTrpLysAsnVal 251
181 AAGGGGTAATTCGCTGGAGCTAGTACAA.....GATCCCAAGAA 224
::: || ||| ::: ||||| ::::: ||| ||
253 PheSer...AsnSerLeuAspGlyValIlyTrLysLeuThrLeuProAlaY 267
223 GGGGGCCATGGCCTTTCAGACTTATGTCACGCTGACCATGCTGCATCGC 274
| :::: ||||| ||::: :::::
267 sValSlYserMetThrAsnSerLysLeuLeuAlaCysGlnLeuLysP 284
275 ACACGCCGCCCA.....ACCAACAAGAGCTAAANAATA 306
::: ||||| ::::: |||||::|||
284 heAlaAlaProLeuLeuAlaLeuLysThrSerThrAspArgLeuSerIle 300
307 ARGGAGCGCTCATTTTACGCGCTCGCTATTGCTGCTGGAGAAATCGCG 356

```

```

301 SerAsnArgSerAsnLeuSerSerAsnMetPrpValMetPheAspGln 317
      :::::::::: :::::::::: :::::::::: ::::::::::
357 AGCAAGCGGCTCGTGAGACGACGAGCATGTACAAATACCGTGGAGAGTGGT 406
      :::::::::: :::::::::: :::::::::: ::::::::::
317 SpPheUSeRAlaThrValValPheProTyrMetHisTyrGlnAsnGlyp 334
      :::::::::: :::::::::: :::::::::: ::::::::::
407 ACAAGTTCATCGAAGATGCCATTACCTG..... 435
      :::::::::: ::::::::::
334 heLeuSerPheSerHisLeuValHisLeuTrpGlySerPheArgAlaSer 350
      ::::::::::
436 CAGCGCGACCTCATATATATCTCGCGACCTGCCGCGAGTGGCGGTACGA 485
      :::::::::: :::::::::: :::::::::: ::::::::::
351 HisGlyAspAsnHisLeuLeuAsnLeuSerGlnGlnAsnLeuGln 367
      :::::::::: :::::::::: :::::::::: ::::::::::
486 AGGCAATCCGCGAGCGGCTGCTTGTCGAGCAAGTGGCTGCCGCTTAAGT 535
      :::::::::: :::::::::: :::::::::: ::::::::::
367 uArGyAllyAspArgAsnArgGlySerGlnGlyLysCysValSerLeuGln 384
      :::::::::: :::::::::: :::::::::: ::::::::::
536 ACCTTCAGGAGCTG.....CATGAGTTCACACAGGATGGTGGT 573
      :::::::::: :::::::::: ::::::::::
384 IsileArgMetLeuAsnSerGlyTyrHisAlaValTyrCysAlaTrpLeu 400
      :::::::::: :::::::::: ::::::::::
574 ATACACCCAGAGA...CGACCGCAG..... 594
      :::::::::: ::::::::::
401 LeuValGlnAsnPheThrPrpGlnGluLeuValGluValCysPheAsnThr 417
      ::::::::::
595 .....TCGTGCAAGCTCTTAAGTCTCGCATCCG 622
      ::::::::::
417 TrpSerGlnIleThrAspLeuArgSerSerLysProSerPheLeuGlnGlysh 434
      :::::::::: :::::::::: ::::::::::
434 IsValSerIleGlnAspVal.....LeuTyrSerSer 444
      ::::::::::
673 ATATTGCAAGCC 684
      ::::::::::
445 IlePheAsnAla 448

```

```

seq_name: sp_bacteria:Q9RXL2
seq_documentation_block:
AC Q9RXL2: PRELIMINARY; PRT; 207 AA.
DT 01-MAY-2000 (TIREMBLrel. 13, Created)
DT 01-MAY-2000 (TIREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TIREMBLrel. 16, Last annotation update)
DE DEOXYGUANOSINE KINASE/DEOXYADENOSINE KINASE SUBUNIT.
GN DR0298.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Hatt D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Uitterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.
RA Kachuma A.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans RI.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001891; AAF09882.1; -.
DR TIGR; DR0298; -.
DR InterPro; IPR002624; -.
DR Pfam; PF01712; dnf; 1.
KW kinase.
SQ SEQUENCE 207 AA; 23873 MW; 3B135D1CFEED0741 CRC64;

```

alignment_scores:

Quality: 148.00 Length: 188
Ratio: 1.345 Gaps: 6
Percent Similarity: 58.511 Percent Identity: 26.064

alignment_block:

US-09-416-579a-1 x Q9RXL2 ..

Align seg 1/1 to: Q9RXL2 from: 1 to: 207

```

67 GTTCCTCATGAGGCGCAACATCGCAGCGGAGACACATATTGAAACA 116
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
3 ValAlaValSerGlyAsnIleGlySerGlyLysSerThrLeuThrArgLe 19
117 CTTC..GAGAGTACAGAGACGACATTGGCTGCTGACCGACCCGCTGC 163
   : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
19 tleuAlaGluArgTyrGly.....LeuArgProValT 30
164 AGAAGTGGGCAACGTCACAGGGGTAATCTGCTGAGCTGATGTACAA 213
   : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
30 TyrIleProTyrAlaGluAsn.....ProTyrLeuGluAspHeterHis 44
214 GATCCCAAGAGTGGGCGCATGCCCTTTCAGAGTTATGTCACGCTGACCAT 263
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
45 AspMetArgGlnTyrSerPheHisSerGlnValTyrPheLeuSerArgAr 61
264 GCTGCAGTGGCACACCCGCC..CCAACCAAGAAAGCTAAAAATTAAGG 310
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
61 gleuGluGlnIleAsnIleuGlyMetValThrGlyAlaArgTyrValIleGln 78
311 AGCGCTCCATTTTTCAGCTGCTATGCTGCTGAGAGAACATGCGACGA 360
   : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
78 spArgThrValPheGluAspAlaAsnIlePheAlaArgAsnLeuTyrGlu 94
361 AACGCTCGCTGGAGCAGGCGCATGTACAATACGCTGAGAGAGTGTACAA 410
   : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
95 SerGlyIleMetGlyIleuArgAspTrpGlnThrTyrArgGlyLeuTyrGln 111
411 GTTCATCGAAGATCCATTCACCTGAGCGGACCTCATCATATATATC 460
   : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
111 uGlyValIleuProAlaLeuArgVal...ProAspLeuIleHisIle 127
461 GCACCTCGCGGAGTGGCGTACGAGACGATCGGCGGCGCTGCTGT 510
   : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
127 spAlaGlyLeuProThrLeuArgArgArgIleAlaLeuArgGlyArgAsp 143
511 GAGGAGAGCTGCGCGCTTAAGTACCTTCAGAGAGCTGCATGAGTTGCA 560
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
144 TyrGluGlnAlaIleProAspGlnTyrLeuAlaGlyLeuAsnArgLeuTyr 160
561 CCGAGCATGCTGTATACACAGAGAGAGACCGAGCTGTCGAAGCTCTG 610
   : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
160 rAlaGlyTyrPile.....AlaAlaPheAspLeuGlyProValVala 174
611 TCCTCGATCCCGCAT 624
   : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
174 rGlyAlaSpIlyasp 178
seq_name: sp_virus:040641
seq_documentation_block:
ID 040641 PRELIMINARY; PRT: 527 AA.
AC 040641:
DT 01-JAN-1998 (TREMBLrel. 05. Created)
DT 01-JAN-1998 (TREMBLrel. 05. Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15. Last annotation update)
DE THYMIDINE KINASE.
OS Salnitiline herpesvirus 2.
OC Viruses: dsDNA viruses, no RNA stage: Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OK NCBI_TaxID=10381;
RN [1]

```

RP SEQUENCE FROM N.A.

RC STRAIN-C-488;
RX MEDLINE=98037620; PubMed=9371569;
RA Knappe A., Hiller C., Thurnau M., Wiltmann S., Hofmann H.,
RA Fleckenstein B., Fickenscher H.;
RT "The superantigen-homologous viral immediate-early gene iel4/vsag in
RT herpesvirus saimiri-transformed human T cells.";
RL J. Virol. 71:9124-9133(1997).
DR EMBL: Y13183; CAI73635.1; ..
DR InterPro: IPR001889; ..
DR Pfam: PF00693; TK_Herpes; 1.
DR ProDom: PD001519; ?; 1.
SQ SOURCE 527 AA; 60040 MW; 44975CC07265F3F2 CRC64;

alignment_scores:

Quality: 136.50 Length: 250
Ratio: 1.075 Gaps: 14
Percent Similarity: 50.800 Percent Identity: 27.200

alignment_block:

US-09-416-579a-1 x 040641 ..

Align seg 1/1 to: 040641 from: 1 to: 527

```

61 TTCACGTCCTCATGAGGCGCAACATCGGCGGAGAACCAAGTATT 110
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
210 PhePheIlePheLeuGluGlySerIleGlyValGlyIleThrThrLeu 226
111 G.....AACCACTCGAAGATACAGAACAGCATTTGCTGCTGA 151
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
226 uLysSerMetAsnGlyIleLeuGlyGlyAsn...ValLeuAlaPhe 242
152 CCGAGCCCGTGGAGAGTGGGCGCAACGTCACAGGCGTAAATTCGTCGAG 201
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
242 IsgIleProIleAlaTyrThrPheAspValPheSer...AsnSerLeuGlu 257
202 CTGATGTACAA.....GATCCCAAGAGTGGGCGCATCCCTTTCAGAGTT 247
   : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
258 GluValTyrIleThrLeuProAlaIleValGlyArgThrSerAsn 274
214 .....GATCCCAAGAGTGGGCGCATCCCTTTCAGAGTT 247
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
274 rAlaIleLeuLeuAlaCysGlnLeuIleuLysPheAlaSerProLeuLeuAla 291
248 ATGTCACGCTGACC..ATGCTGCAGTCGCACACCGGCCCAACCAAG 294
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
291 euLysThrAlaThrAspArgLeuSerSerHis.....LysAsnSer 304
295 AAGCAAAA.....ATAAGGAGCGCTCCATTTTTCAGCGC 329
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
305 LeuLeuSerSerAspMetTrpValMetPheAspArgHisProLeuSerAl 321
330 TCGCTATTGCTGTCGAGAACATCGACGAAACGCTGCTGAGAGCAG 379
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
321 aThrValValPheProTyrMetHisPheGlnAsnGlyPheLeu.....S 336
380 GCATGTACAATACGCTGAGAGGTACAGATTCACGAGAGTCCAT 429
   : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
336 erPheSerHisLeuIleGlnLeuTrpSerPheLysAlaSer..... 350
430 CACCTGCAGGCGGACCTCATATATCTGGCAGCTCGCGGAGGTGGC 479
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
351 .....GlnGlyAspAsnIleIleLeuLeuAsnLeuAsnSerGlnGluAs 365
480 GTACAGACGATCGGCGGCGCTGCTTCGAGGAGACGCTGCTGCCGC 529
   : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
365 nLeuLysArgValArgLysArgAsnArgLysGlnGlnLysSerValSer 382
530 TTAGTACTTCAGAGAGCTG.....CATGAGTTCACAGAGAC 567
   : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
382 IeGluHisIleArgLeuLeuAsnAsnCysTyrHisAlaValTyrCysAla 398

```


OS Myxococcus xanthus.
 OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
 MX Myxococcales; Cystobacteriales; Myxococcaceae; Myxococcus.
 RX NCBI_TaxID=34;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DK 101;
 RA Harris B.2., Singer M.H.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF055904; AAC82365.1;
 KW Hypothetical protein
 SQ SEQUENCE 542 AA; 57203 MW; DA7E39075A86A8BC CRC64;

alignment_scores:

Quality: 135.00 Length: 272
 Ratio: 1.107 Gaps: 19
 Percent Similarity: 44.853 Percent Identity: 30.515

alignment_block:

US-09-416-579a-1 x 068872 ..

Align seg 1/1 to: 068872 from: 1 to: 542

```

14 CATCTGTCGCCGAAAG.....GACCAAGTACGCCGAGGACCCAG 57
   ::::: |||::: ||||| ::::: ||||| |||||
44 TYRPROProGlnAlaValProGlyProCysAlaProAlaProPr 60
   ::::: |||::: ||||| ::::: ||||| |||||
58 CCCTTCACGCTCATGAGGCGCAACGCGGACGAGGAGACCACTA 107
   ||||| ||| ||| ::::: ||||| ::::: |||||
60 oargSerPro.ProProAlaGlyAlaGlySerAlaGlySer 76
   ||||| ||| ||| ::::: ||||| ::::: |||||
108 TTGGAACCACTTGAGAACTACAGACGACATTTGCTGTGACGAGC 157
   ||::: |||::: |||::: |||::: |||::: |||:::
77 PheArgPro..ThrcysAlaAlaGlySerGlyAlaGlyCys...ProAla 91
   ||||| ||| ::::: ||||| ::::: |||||
158 CCGTGAGAGAGTGCGGCAACGTCACAGGGGTAATCTGCTGAGCTGATG 207
   ||||| ||| ::::: ||||| ::::: |||||
92 ProSerAlaGlyAlaAlaGlyAlaGlyPro..... 100
   ||||| ||| ::::: ||||| ::::: |||||
208 TACAAGATCCCAAGAGTGAGCA.....TGC...CCTTCAGAG 245
   ::::: ||||| ||| ||| ||| |||
101 .....AlaGlyProSerAlaGlyCysGlyProProAlaGly 112
   ::::: ||||| ||||| ||||| ||||| |||||
246 TTATGTACGCTGACCATGCTGCACTGCACACCG...CCCCAACACAA 292
   ::::: ||||| ||||| ||||| ||||| |||||
112 ySerThrProThrAlaGlyCys.....ThrProCysProAlaProPro 126
   ||||| ||| ||| ||| ||| |||
293 AGAAGCTAAATATAGAGGCGCTCATTTTACGCGCTGATTCCTTC 342
   ||||| ||| ||||| ||||| ||||| |||||
127 ATGcysAlaGlyAlaAlaGlyProSerAlaPro.....AlaProSerAl 139
   ||||| ||| ||||| ||||| ||||| |||||
343 GTGAGAACATGCGAGCAACGCTGCTGAGAGGAGCA.....TGTA 386
   ||||| ||| ::::: ||||| ||||| ||||| |||||
139 aglyArgThrCysProThrAlaGlyProSerCysAlaSerGlyCys 156
   ||||| ||| ::::: ||||| ||||| ||||| |||||
387 CAATACGCTGAGAGTGATCAAGT..... 412
   ||||| ||| ::::: ||||| ||||| ||||| |||||
156 ysProThrGlyAlaGlySerAlaProGlyProThrProHisGlyAla 172
   ||||| ||| ::::: ||||| ||||| ||||| |||||
413 .....TCATGAGAGTCACTCACTGAGGAGGAGGAGGAGGAGGAGGAG 450
   ||||| ||| ::::: ||||| ||||| ||||| |||||
173 GluProSerProProSerGlnSerPro.SerProAlaProAlaSer... 187
   ||||| ||| ::::: ||||| ||||| ||||| |||||
451 ATATATCTGCGACCTCGCGAGAGTGCGGTACGACGACGACGAGCG 500
   ::::: ||||| ||||| ||||| ||||| |||||
188 .....SerGlyAlaGlyAlaGlyThrHisProArgPro 198
   ||||| ||| ::::: ||||| ||||| ||||| |||||
501 GGCTGCTTCTAGAGAGAGCTGCGCTTAAGTACCTTACGAGAGCTGC 550
   ||||| ||| ::::: ||||| ||||| ||||| |||||
199 Ala.....ArgAla.....SerAlaAlaG1 205
   ||||| ||| ::::: ||||| ||||| ||||| |||||
551 ATGAGTTGCACGAGACTGTTGATACACGAGAGACGAGCGAGTCTGC 600

```

seq_name: sp_bacteria:051255

seq_documentation_block:

ID 051255 PRELIMINARY; PRT; 205 AA.
 AC 051255;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE DEOXYGUANOSINE/DEOXYADENOSINE KINASE(1) SUBUNIT 2 (DCK).
 GN BB0239.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
 RA Peterson J., Kierlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA Van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,
 RA Ulterback T., Wathey L., McDonald L., Artlach P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi";
 RL Nature 390:580-586(1997).
 DR EMBL: AE001134; AAC66630.1; -;
 DR TIGR: BB0239; -;
 DR InterPro: IPR002624; -;
 DR Pfam: PF01712; dNK; 1.
 DR PIR: P01712; dNK; 1.
 SQ SEQUENCE 205 AA; 24116 MW; F25F410DD7F1C6F CRC64;

alignment_scores:

Quality: 132.50 Length: 196
 Ratio: 1.380 Gaps: 9
 Percent Similarity: 48.980 Percent Identity: 24.490

alignment_block:

US-09-416-579a-1 x 051255 ..

Align seg 1/1 to: 051255 from: 1 to: 205

```

67 CTCCTCATGAGGCAACATCGGACGCGGAGAGACGATATTGAC.. 114
   ::::: ||||| ||||| ||||| ||||| |||||
7 TleValIleGlyLeuIleGlyValGlyLysThrThleuGlyAsn1 23
   ||||| ||| ::::: ||||| ||||| ||||| |||||
115 .....CACTCGAAGTACAGAAAGACGACA 139
   ::::: ||||| ||||| ||||| ||||| |||||
23 eleuSerLysGluLeuGluValProPheTyrSerGluLeuAsnAsp 40
   ||||| ||| ::::: ||||| ||||| ||||| |||||
140 TTTCCTGCTGACCGAGCCGTCGAGAAAGTGGCGCAACGTCAACGGGTA 189

```

```

40 heThLeuAla..... 43
190 AATGCTGAGCGTATGACAAAGATCCCAAGAAAGTGGCCATGCCCTT 239
44 ...ValLeuAspLysPheTyrLysAspLysSerArgTPrAlaPheProVa 59
240 TCAGAGTTATGTCACGCTGACATGCTGACAGTGCACACCGCCCAACCA 289
59 IGLn.....IleAsnPheLeu.....A 65
290 ACAGAAGAGCTAAATATAGAGACGCGTCCATTTTASCGCTGCTATTCG 339
65 snGluArgPheLysLeuIle...LysGlyValPheArgTyr..... 77
340 TTCCTGGGAACATGCGCAAGCGCTGCTGAGCAGGCGCATGTAC... 387
78 .....LysGlyGlyIleLeuAspArgSerIleTyrG1 88
388 .....AATACCTGGAG 400
88 YAspCysValPheAlaSerLeuLeuAsnCysAspGlyHisIleSerAspG 105
401 AGTGTACAACTCATGAGAGATCCAT.....CACTG 435
105 LngLutTyrLysIleTyrIleAspLeuAspAsnMetLeuGlnHisSer 121
436 CAG...GGGAGCTCATCATATATATCTCGCACCTCGCCGAGGTGGCTA 482
122 GlnArgProSerIleLeuValTyrLeuAspCysSerIleAspGlnValG1 138
483 CGAAGCCTCCGCGACGCGGCTCTGTCTGAGAGAGCTGCGTCCGCTTA 532
138 nArgArgIleLysAsnArgAsnArgSerPheGlnMetAsnIleProArg 155
533 AGTACCTTCAGAGCTGCATGATTCGACACGAGACTGG 570
155 sPTyrLeuGlnGlyLeuAsnArgLysTyrIleuLysTrrp 167

seq_name: sp_plant:Q9M715

seq_documentation_block:
ID Q9M715 PRELIMINARY; PRT; 274 AA.
AC Q9M715;
DT 01-OCT-2000 (Trembl) 15, Created)
DT 01-OCT-2000 (Trembl) 15, Last sequence update)
DT 01-MAR-2001 (Trembl) 16, Last annotation update)
DE ARABINOGALACTAN PROTEIN (FRAGMENT).
GN AGP.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Ponce G., Lujan R., Campos M.E., Reyes A., Nieto-Sotelo J.,
RA Feldman L.J., Cassab G.I.:
RT "The quiescent center controls the expression of three cap genes in
RT the maize root meristem."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF134579; AAF3497.1; -.
DR InterPro: IPR002965; -.
DR PRINTS: PR01217; PRICHEXTNSN.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 274 AA; 27957 MW; BFE373BB6CA8B713 CRC64;

```

alignment_scores:

Quality:	132.00	Length:	259
Ratio:	1.128	Gaps:	16
Percent Similarity:	45.174	Percent Identity:	30.116

alignment_block:

US-09-416-579a-1/rev x Q9M715 ..

Align seg 1/1 to: Q9M715 from: 1 to: 274

```

745 CGAGCCCTGGCGCTGCTGGCGGCACACAGAACCGGCGGCTGTGG 696
7 ArgAlaSerSerAlaIleThrAlaThrProSerThrSerThrAlaIleG1 23
695 TTACTTGAGATGGCGTGAATATGCTCTCCG..... 662
23 Y.....ThrProThrSerLacysSerProThrAlaThrSerThrs 37
661 .....AGCGTGTACTCGTGGCAATGTTTCCAGGTCA 626
37 eThrArgThrSerSerAlaSerThrAlaProThrAlaCysProGlyThr 53
625 GATCGCATGAGGAGCTAGAGCTTCGACGACACTGCGTCTGTGTTG 576
54 SerArgGlySerArgProSerProCysSerSerThrAlaThrSerSerTh 70
575 ATCAACCACTCTGCTGTCACACTCATGCACTCTGAAGT.....ACTT 532
70 rSer...AlaProGly.....ArgArgProArgGlyThrThrThrs 83
531 AAGCGGACGCGAGCTCTCCAGAAC...GAGCGCGTGGCGGATCGTT 485
83 eThrAlaThrPserSerProThrAlaSerProCysAlaSerCys... 98
484 CGTACGCGACCTCCGCGGAGGTGGCAGATATATGATGAGTCCGCTTCG 435
99 ArgGlnProThrProArg.....GlyArgProAla 108
434 AGCTAATGACTCTTCGATTAACCTTACACACTCCACAGGATATGTA 385
108 A.....ProCysArgArgCysP 114
384 CATGCCCTGCTTCACAGCGCGTTTCGTCGAGTTCACAGAACAT 335
114 rSerProAlaProAla...ArgProThrAlaCysSerSerArgSerThr 129
334 AGCGAGCGCTAAATAATGAGCGCTCTTATTTTACTTCTTGTGTT 285
130 Ala.....GlySerProSerGlyProThrProCys..... 139
284 GGGCGCGTGTGAGATGACGATGTCAGCGTGACATTAACGTGAAAGG 235
140 .....ProSerGlnArg 144
234 CATGGCCACTTCCTGGGATCTTTGTACA.....TCAGTCCAGCA 194
144 rSerProGly.....CysThrGlyThrAlaSerArgProThr 156
193 GATTACCCCGTTGACGTTGCGCCCACTTCGACGGGTCTCGTACGAG 144
157 ThrAlaSerArgThrLeuThrTrpProSerSerSerAlaArgSerProp 173
143 CAATGCTGTTCTGTACTTCGAAAGTGTTCAATATCGTGTCT... 98
173 oThrCysThrAlaThrPserAlaArgThrAlaProThrThrSerThrG 190
97 .....TCCCGTGGCGAGTGGCTCTCGATGAGAGCGTGAAGGCT 56
190 LysMetSerArgProGlnCysProProThrGlyGly.....ThrAla 204
55 GGGTGGCTTGGCGCTACTGTCCTCT 29
205 ThrThrProArgProAlaSerSerPro 213

seq_name: sp_virus:Q83342
seq_documentation_block:

```


RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U97553; AAB6394.1; ..
 DR EMBL: AF105037; AAF19286.1; ..
 DR InterPro: IPR001889; ..
 DR Pfam: PF00693; TK_L herpes; 1.
 DR ProDom: PD001519; ..
 KW Kinase.
 SQ SEQUENCE 644 AA; 72269 MW; FD3CD1646FE5562 CRC64;

alignment_scores:

Quality: 132.00 Length: 275
 Ratio: 0.957 Gaps: 11
 Percent Similarity: 50.182 Percent Identity: 22.909

alignment_block:

US-09-416-579A-1 x 041940 ..

Align seg 1/1 to: 041940 from: 1 to: 644

```

22 GCCCGAAGGGGACCAAGTACGCCGAGGC..ACCCAGCCCTTCACCGT 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
312 SerlysmetSerThrAspPheProAlaGlyProThrArgCysAlaThrVa 328
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
69 CCGC..ATCGAGGGGCAACATCGCAGCGGAGAGACACAGCTATTGAACC 115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
328 LLeuThrPheAspGlyLeuAlaValGlyLysThrThrAlaLeuGlnA 345
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
116 ACTTCGAGAACTACAGACACACATTTGGCTGCTGACCGAGCCCGTCGAG 165
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
345 LaAlaAlaGluPheLeuHisAsnLeuValLLeuSerGluProLLeuPro 361
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
166 AAGTGGCCGCAACGTCACGCGGGTAATCTGCTGGAGCTGATGATC.... 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
362 TyrTrpThrSerThrPheAspLysAsnValCysGlnGlnIleTyrAspVa 378
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
211 .....AAAGATCCCAAGAG..... 225
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
378 LValLysThrLysLysLysGlyLysLysHisSerLysLysValLeuGlnC 395
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
226 .....TGGCCATGCCCTT.....CAGATGATGTCACG 255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
395 ysglmetAlaPheAlaGlnProPheHisAlaThrGlnHisLeuLeuArg 411
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
256 CTGACCATGCTGCAGTCGACACACCGCCCA..ACCAACAGAGCTAAA 302
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
412 ArgThrCysLeuGlyThrLysValSerAspGlyThrCysAsnAsnTyrVa 428
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
303 AATAAGGAGCGCTCATTTTACGCGCTGATGCTGCTGAGAGACA 352
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
428 LLeuIleAspArgHisMetValSerProThrValLLeuPheProCysLeuP 445
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
353 TGGCAGCAAGAGCGCTCGCTGAGCAGGAGCATGTACATAGCCTGAGAG 402
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
445 hePheAlaGlyValLeuArg.....PheCysAspLeuLeuSerLeu 459
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
403 TGCTACAGATTTCATCAAGAGAGTCATACCTCAGCGAGCATCATCAT 452
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
460 LeuSerValPheSerAlaGlnLysPhe.....AspAsnValVa 472
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
453 ATATCTGGCAGCTCGCCGAGAGTGCGTACGAACGATCCGCGAGCGGG 502
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
472 LLeuPheLysLeuHisProLysValAlaAlaThrArgValLysGlnArg 489
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
503 CTGCTGTGAGAGAGCTGGCGCTTAAAGTACCTTCAGAGAGCTCAT 552
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
489 LArgAsnProGluGlyValLLeuHisThrThrLysGlnLeuLeuAsn 505
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
553 GAG.....TGACACAGAGAGCTGGTGTACACACAGAGACAGC 590
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
506 GluValLeuAspAlaLLeuPheCysAlaLTrpLLeuLeuAlaLLeuLTr 522
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
591 GCACTGCTGCAAGCTCTAGTCTCGATGCCGATCTGACCTGGAAAAA 640

```

seq_name: sp_human:09H7T3

seq_documentation_block:

ID 09H7T3 PRELIMINARY; PRT; 257 AA.
 AC 09H7T3;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE CDNA FLJ14280 F1S, CLONE PLACE1005584, WEAKLY SIMILAR TO TRANS-ACTING
 DE TRANSCRIPTIONAL PROTEIN ICP0.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA WagaSuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakematsu A., Nakamura Y., Nagahari K., Masuno Y.,
 RA Niomiya K., Iwayanagi T.,
 RT "NEBO human cDNA sequencing project."
 RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK024342; BAF14892.1; ..
 SQ SEQUENCE 257 AA; 26269 MW; B914A57EA54DF418 CRC64;

alignment_scores:

Quality: 129.50 Length: 258
 Ratio: 1.295 Gaps: 18
 Percent Similarity: 38.760 Percent Identity: 27.519

alignment_block:

US-09-416-579A-1/rev x 09H7T3 ..

Align seg 1/1 to: 09H7T3 from: 1 to: 257

```

730 TGCTGGCGCAGACAGACCGGAGGCGTGTGTTACTTGAATGGCG 681
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
49 CysThrAlaGlyArgArgProAsnArgAlaSerGly.....ArgArgArg 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
680 TCGAATATGCTGC...TTCGAGCGCTGTTACTCGTCCCAATGTTTTC 634
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63 galsercysSerProAlaProThrTrpProLeucys...CysTyrP 79
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
633 CAGGTCACAGTCGGCATCGAGAGACTAGACCTTCACAGACCTCGGCTGC 584
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
79 roGlnSerArg.....ProThrAlaSerAla 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
583 TCTGGTGTATCAACAGTCTGTCGCACTACGAGCT.....CC 543
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
88 AlaGly.....ProGlyAlaCysMetArgAlaSerGlyArgPr 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
542 TGAAGTACTTAAACGCGACGAGCTTCCTCAGAACGAGCCC..... 500
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
100 OHISGLYAsnThrThrAlaSerThrAlaProAlaArgHisProArgProA 117
499 ....GCTGCCGATGCGTTCGACCCACCTCCGGCGAGTGGCGAGAT 455
      ||||| ||| ||| |||||
117 rGArgProSlyGLyProAlaLeuArgProThrProArgProCysAla... 132
454 AATATGATGAGTCCGCTGCAGTGAATGACTCTTCATGAACCTGTATC 405
      |||||
133 .....GLyPro.....Al 135
404 CACTCCTCCAGCGTAT.....TGATACGCCCTGCTCCA..... 371
      : ||||| ||| |||
135 aProProAlaSerArgAspCysArgCysArgProArgArgTrpP 152
370 .....GCGAGCCGTTGTCGCATGTCTCCACGAGAC 338
152 roArgAlaGLyArgArgGLyArgArgAlaGLyAlaCysLysPro..... 166
337 AATAGCGAGCGCTAAATAATGAGCGCTCCTTATTTTAGCTTCTGTG 288
166 ..... 166
287 GTTGGGCGGTGTCGACTGCAGCATGTCCAGCGTGACATACTTGAA 238
      ||||| ||||| |||||
167 .....SerCysAlaGLyAlaAlaTrpSerAla.....A 176
237 GGGCATGGCCCACTTCTTGGAATCTTGTACATCAGCTCCAGCAGATT 188
176 rGLyAlaPro.....LeuCys..... 181
187 CCCGTTGAGTTCGCGCACTTCTCGACGGGCTCGGTCCAGCAGCAATG 138
      ||| ||||| |||||
182 .....SerTyrArgThrSerCysAlaGLySerC 191
137 TCGTTCCTTGACTTCGAAGTGGTCAATAACGTGCTCTCC..... 95
      | :||| :||| :|||
191 sGLyAlaArgThrAlaProThrProAlaProThrCysAlaSerProSera 208
94 .....CGCTGCCGATGTGCCCTCATGAGAGAGCGTGAAGGCGCTGGTGC 50
      ||||| ||| |||
208 laAlaAlaSerSerCysCysArgArg.....ArgArgAlaCys 220
49 .....CCTCGGCGTACTTGCTC 32
      ||| :|||
221 SerSerProThrThrAlaTrpSer 228
```


Quality: 492.00 Length: 215
Ratio: 2.911 Gaps: 3
Percent Similarity: 78.605 Percent Identity: 45.116

alignment_block:

US-09-416-579a-1 x KITM_MOUSE ..

Align seg 1/1 to: KITM_MOUSE from: 1 to: 270

```

67 GTCCTCATGAGGCAACATCGGACGAGGAGACGATTTGAACCA 116
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
58 ValCysIleGluGlnIleAlaSerGlyLysThrCysLeuGluH 74
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
117 CTTCGAGAGTACAGACGATTTGCTGCTGACCGCCGCTGAGA 166
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
74 epheserA...ThrThsAlaValGluLeuMetGluProValLeu 90
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167 AGTGGGCAACGTCACAGGGGTAATCTGCTGAGCTGATACAAAGT 216
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
90 ystrpAgnValHisGlyHisAsnProLeuSerLeuMetLysHisAsp 106
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
217 CCCAGAGATGGGCAATGCTTTCAGATTAATGTCAGCTGACCATGCT 266
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
107 AlaserAtrpGlyLeuThrLeuGlnTrpValGlnLeuThrMetLe 123
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
267 GCAGTGGCAGACGCGCCCAACCAAGAGCTAAATATAGAGAGCGCT 316
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
123 unspGlnHisThrArgProGlnMetSerProValArgLeuMetGluArg 140
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
317 CCATTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
140 eTlleTyrAlaArgTyrIlePheValGlnAsnLeuTyrArgGlyGly 156
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
367 TCGCTGGACAGGCGCATGACATACGCTGAGAGAGTGTACAGTGCAT 416
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
157 LysMetProGluValAspTyrAlaIleLeuSerGluTrpPheAspTrp 173
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
417 CGAAGAGTGCATTCACCTGACGCGGAGCTCATATATCTGCGACCT 466
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
173 eValAlaGlnIleAspValSerValAspLeuIleValTyrLeuArgTrp 190
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
467 CGCCGAGAGTGGCGTACGAGACGATCGGACGCGGCTGCTGCTGAGAG 516
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
190 hrProGluIleCysTyrGlnArgLeuLysMetArgCysArgGluGlu 206
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
517 ACTGCGTCGCGCTTAAGTACCTTACAGAGCTGATGATGATGACAGCA 566
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
207 LysValIlePheMetGluTyrLeuHisAlaIleHisArgLeuTyrGlu 223
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
567 CTGGTTGATACAC...CAGAGACGACGCGCATGCGCAAGTCCATGCC 613
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
223 utrpLeuValAsnGlySerLeuPheProAlaAlaProValLeuVal 240
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
614 TCGATGCGCATCTGAACCTGGAACATTTGAGCAGAGTACAGGCTGCG 663
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
240 LeGluAlaAspHisAsnLeuGluLysMetLeuGluLeuPheGluGln 256
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
664 GAGACGACGATTTGAGCGCATCTCAAGTACCAACAGCCCT 706
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
257 ArgAlaArgIleLeuThrProGluAsnTrpLysHisGlyPro 270
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

seq_name: SwissProt_39:KITM_HUMAN

seq_documentation_block:

ID KITM_HUMAN STANDARD; PRT; 232 AA.
AC 000142; 015238;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE THYMIDINE KINASE 2, MITOCHONDRIAL (EC 2.7.1.21) (MT-TK).
GN TK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97236800; PubMed=9079672;
RA Johansson M., Karlsson A.;
RT "Cloning of the cDNA and chromosome localization of the gene for
human thymidine kinase 2."
RL J. Biol. Chem. 272:8454-8458(1997).
RN [2]
RP SEQUENCE FROM N.A., SEQUENCE OF 1-28, AND CHARACTERIZATION.
RC TISSUE=Brain;
RX MEDLINE=99142705; PubMed=9989599;
RA Wang L., Munch-Petersen B., Herrstroem Sjoelberg A., Hellman U.,
Bergman T., Joernvall H., Eriksson S.;
RT "Human thymidine kinase 2: molecular cloning and characterisation of
the enzyme activity with antiviral and cytostatic nucleoside
substrates."
RL FEBS Lett. 443:170-174(1999).
CC -1- FUNCTION: DEOXYRIBONUCLEOSIDE KINASE THAT PHOSPHORYLATES
THYMIDINE, DEOXYCYTIDINE, AND DEOXYURIDINE. ALSO PHOSPHORYLATES
ANTI-VIRAL AND ANTI-CANCER NUCLEOSIDE ANALOGS.
CC -1- CATALYTIC ACTIVITY: ATP + THYMIDINE = ADP + THYMIDINE
5'-PHOSPHATE.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN LIVER, PANCREAS,
MUSCLE, AND BRAIN.
CC -1- SIMILARITY: BELONGS TO THE DCK/DCK FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb.ch/announce/
or send an email to license@isb.sib.ch).
CC EMBL: U77088; AAC51167.1; -
CC EMBL: Y10498; CAA71523.2; -
CC MIM: 188250; -
CC DR InterPro: IPR002624; -
CC DR Pfam: PF01712; DCK; 1.
CC KW Transferase; Kinase; DNA synthesis; ATP-binding; Mitochondrion.
CC FT NP_BIND 24 31 ATP (PORENTIAL).
CC FT CONFLICT 1 8 VORYAMP -> MGAFQGRPS (IN REF. 1).
CC FT CONFLICT 28 28 S -> G (IN REF. 1).
CC SQ SEQUENCE 232 AA; 27494 MW; 479C66CB57F1AC0F CRC64;

```

alignment_scores:
Quality: 488.00 Length: 225
Ratio: 2.789 Gaps: 3
Percent Similarity: 77.778 Percent Identity: 43.111

alignment_block:

US-09-416-579a-1 x KITM_HUMAN ..

Align seg 1/1 to: KITM_HUMAN from: 1 to: 232

```

37 AGTACGCGGAGGCAACCGCCCTTACCGTCTGATGAGGCAACAT 86
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
10 LysGluGlnIleLysGlnLysSerValIleCysValGluGlnLys 26
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
87 CGGACGCGGAGGCAACCGATTTGACCACTTGCAGAGTACAGAGC 136
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
26 eAlaSerGlyLysThrCysLeuGlnPhePheSerAsn...AlaThr 42
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
137 ACATTTGCTGCTGACGAGCGCGTGCAGAGAGGCGCAACGCAAGG 186
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
42 spValGluValLeuThrGlnProValSerLysTrpArgAsnValArg 58
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

RX MEDLINE=91192170; PubMed=2013338;
RA Eriksson S., Cederlund E., Bergman T., Joernvall H., Bohman C.;
RT "Characterization of human deoxycytidine kinase. Correlation with
RT CDNA sequences.";
RL FEBS Lett. 280:363-366(1991).
RN [3]
RP SUBCELLULAR LOCATION.
RX MEDLINE=98004502; PubMed=9342341;
RA Johansson M., Brismar S., Karlsson A.;
RT "Human deoxycytidine kinase is located in the cell nucleus.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:11941-11945(1997).
CC -1- FUNCTION: REQUIRED FOR THE PHOSPHORYLATION OF SEVERAL
CC DEOXYRIBONUCLEOSIDES AND CERTAIN NUCLEOSIDE ANALOGS WIDELY
CC EMPLOYED AS ANTIVIRAL AND CHEMOTHERAPEUTIC AGENTS.
CC -1- CATALYTIC ACTIVITY: NTP + DEOXYCYTIDINE -> NDP + CMP.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE DCK/DCK FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M60527; AAA35752.1; -
DR PIR: A38585; A38585.
DR PIR: S14321; S14321.
DR MIM: 125450; -
DR InterPro: IPR002624; -
DR Pfam: PF01712; dmk.1.
KW Transferase; Kinase; ATP-binding; Nuclear protein.
-FT NP_BIND 28 35 ATP (PROBABLE).
SQ SEQUENCE 260 AA; 30518 MW; 626B9D2D6BED8DBC CRC64;

```

```

-- Alignment_scores:
Quality: 356.50 Length: 247
Ratio: 2.214 Gaps: 5
Percent Similarity: 65.182 Percent Identity: 30.769
Alignment_block:
US-09-416-579a-1 x DCK_HUMAN ..

```

```
Align seg 1/1 to: DCK_HUMAN from: 1 to: 260
```

```

43 GCCGAGGACCCAGCCCTTCACCGCTCTCATCGAGGCAACATCGGACG 92
   :::::::::::::::::::: ::::::::::::::::::::
16 SerGluGlyThrArgIleLysLysLysLysLysLysLysLysLysLys 32
   :::::::::::::::::::: ::::::::::::::::::::
93 CGGGAAGACCGATTTGAACCACTTCGAGAGTACAGACGACATTT 142
   :::::::::::::::::::: ::::::::::::::::::::
32 aglyLysSerThrPheValAsnIleLeuLysGlnLeuLysGlnLys 49
   :::::::::::::::::::: ::::::::::::::::::::
143 GCCGCTGACGAGCCGCTGAGAGAGGCGGACAGTC..... 180
   :::::::::::::::::::: ::::::::::::::::::::
49 LuValValProGluProValAlaArgTyrPcysAsnValGlnSerThrGln 65
   :::::::::::::::::::: ::::::::::::::::::::
181 ..... AACGGCGTAATTCGCT 197
   :::::::::::::::::::: ::::::::::::::::::::
66 AspGluPheGluGluLeuThrMetSerGlnLysAsnGlyGlnValLe 82
   :::::::::::::::::::: ::::::::::::::::::::
198 CGAGCTGATGTACAAAGATCCCAAGAGTGGCATGCCCTTTAGAGTT 247
   :::::::::::::::::::: ::::::::::::::::::::
82 uGlnMetMetLysGluLysProGluArgTyrSerPheThrPheGlnThr 99
   :::::::::::::::::::: ::::::::::::::::::::
248 ATGTACACGACCATGCTGACGAGTGCAGACCGGCGCAACCAACAGAG 297
   :::::::::::::::::::: ::::::::::::::::::::
99 YrAlaCysLeuSerArgIleArgAlaGlnLeuAlaSerLeuAsnGlyLys 115
   :::::::::::::::::::: ::::::::::::::::::::
298 CTAAAA..... ATAAAGGAGCGCTCCATTCTTAG 326

```

```

116 LeuLysAspAlaGluLysProValLeuPhePheGluArgSerValTyrLys 132
   :::::::::::::::::::: ::::::::::::::::::::
327 CGCTCCCTATTGCTTGTGGAGAACATCGGACGAAGCGCTGCTGGACG 376
   :::::::::::::::::::: ::::::::::::::::::::
132 rAspArgTyrIlePheAlaSerAsnLeuTyrGlnSerLysCysMetAsnG 149
   :::::::::::::::::::: ::::::::::::::::::::
377 AGGCGATGTACATACGCTGAGAGAGTGTACAGATTCATCCAGAGTCC 426
   :::::::::::::::::::: ::::::::::::::::::::
149 LuThrGluThrPheIleTyrGlnAspTyrPheAspTyrMetAsnGln 165
   :::::::::::::::::::: ::::::::::::::::::::
427 ATP..... CACCTGAGCGGACGACCTCATATATATCTGCGACCTGCG 470
   :::::::::::::::::::: ::::::::::::::::::::
166 PheGluGlnSerLeuLeuLysAspGlyIleIleTyrLeuGlnAlaThrPr 182
   :::::::::::::::::::: ::::::::::::::::::::
471 GAGAGTGGCGTACGAACGATCCGCGAGCGGCGCTGTTCGAGAGAGCT 520
   :::::::::::::::::::: ::::::::::::::::::::
182 oGluThrCysLeuGlnHisArgIleTyrLeuArgGlyArgAsnGluGlnG 199
   :::::::::::::::::::: ::::::::::::::::::::
521 GCGTGGCGGTTAAGTACCTTCAGAGCTGATGAGTTCAGACGAGAGTGG 570
   :::::::::::::::::::: ::::::::::::::::::::
199 LylleProLeuGluTyrLeuGluLysLeuHisTyrLysHisGluSerThr 215
   :::::::::::::::::::: ::::::::::::::::::::
571 TTGATACACGAGACGAGACG..... CAGTCGTGCA 602
   :::::::::::::::::::: ::::::::::::::::::::
216 LeuLeuHisArgThrLeuLysThrAsnPheAspTyrLeuGlnGluValPr 232
   :::::::::::::::::::: ::::::::::::::::::::
603 GGTCTAGTCTGCTGATGCGGATCTGACCTGGAAGAATGGCAGCAGT 652
   :::::::::::::::::::: ::::::::::::::::::::
232 oLleuThrLeuAspValAsn..... GluAspPheLysAspLysGT 246
   :::::::::::::::::::: ::::::::::::::::::::
653 ACCAGCGCTCGGAGACGACATATTCGAGCGCATTCACAGT 693
   :::::::::::::::::::: ::::::::::::::::::::
246 YrGluSerLeuValGluLysValLysGluPheLeuSerThr 259
   :::::::::::::::::::: ::::::::::::::::::::
seq_name: swissprot_39:DCK_RAT

```

```
seq_documentation_block:
```

```

ID DCK_RAT STANDARD: PRT: 260 AA.
AC P48769;
DR 01-FEB-1996 (Rel. 33, Created)
DR 01-FEB-1996 (Rel. 33, Last sequence update)
DR 01-OCT-2000 (Rel. 40, Last annotation update)
DE DEOXYCYTIDINE KINASE (EC 2.7.1.74) (DCK).
GN DCK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95121928; PubMed=7821805;
RA Stegmann A.P., Hondeers M.W., Willems R., Landegent J.E.;
RT "Cloning of the Dck gene encoding rat deoxycytidine kinase.";
RL Gene 150:351-354(1994).
CC -1- FUNCTION: REQUIRED FOR THE PHOSPHORYLATION OF SEVERAL
CC DEOXYRIBONUCLEOSIDES AND CERTAIN NUCLEOSIDE ANALOGS WIDELY
CC EMPLOYED AS ANTIVIRAL AND CHEMOTHERAPEUTIC AGENTS.
CC -1- CATALYTIC ACTIVITY: NTP + DEOXYCYTIDINE -> NDP + CMP.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE DCK/DCK FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L33894; AAA65098.1; -
DR InterPro: IPR002624; -

```



```

49 GGCACCCAGCCCTTTACCGCTCATCGAGGACATCGCGGAA 98
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
35 GLYArgGlyProArgArgLeuSerIleGluGlyAsnIleAlaValGly 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
99 GACCACGATTTTGAACCTTGGAGATGACAAAGACGATTTGGCTGC 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 sSerThrPheValIleLeuLeuThrIleThrIleProGluIleValAla 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
149 TGACCGAGCCGCTCGAGAGTGGCGCAAGTCAAGCGGTA..... 189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 lathIleProValAlaIleThrIleGlnAsnIleGlnAlaIleGlnGln 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
190 .....AATCTGTGAGCTGATGATACA 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
85 LysAlaCysThrAlaGlnSerLeuGlnAsnLeuLeuAspMetCysIle 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
213 AATCCCAAGAGTGGCCATCGCTTTCAGATGATATGTCACGCTGACCA 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
101 gGIuProAlaArgTyrSerTyrThrPheGlnThrPheSerPheLeuSer 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
263 TGTCTGAGTGCACACCGCCCAACCAACAGAGCTA..... 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
118 rIleLeuIleValGlnLeuGluProPheProGluIleValLeuGlnAlaArg 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 .....AAATAAGGAGCGCTCATTTTACGCTGCTATTGCTT 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
135 LysProValGlnIlePheGluArgSerValTyrSerAspArgTyrIlePhe 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
342 CGTGGAGAACATGCGAGCAAGCGCTGCTGAGCAGGCGCATGTACAATA 391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
151 eAlaLysAsnLeuPheGlnAsnGlySerLeuSerAspIleGluIlePheIle 168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
392 CGCTGAGAGAGTGTACAAGTTCATC.....GAAGACTCATTCACCTG 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
168 lEtyGlnAspTyrPheIleSerPheLeuLeuTyrGluPheAlaSerArgIle 184
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
436 CAGGGGACCTCATATATATGCGACCTCGCCGAGGAGTGGCGTACGA 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
185 ThrLeuHisGlyPheIleTyrLeuGlnAlaSerProGlnValCysLeuIly 201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
486 ACCGATCCGCGAGCGGCTGCTGAGAGAGTGGCTGCGCTTAAAGT 535
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
201 sArgLeuTyrGlnArgAlaArgGluGluGlyIleGluLeuAlaIle 218
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
536 ACCTTCAGAGAGCTGATGATGTCACGAGAGTGTGATACACAGAGA 585
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
218 rIleGluGlnIleLeuHisGlyIleGlnIleAlaTyrPheIleHisIleThr 234
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
586 CGACCG.....CAGTCGTGCAAGTCTCTAGTCTGCA 617
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
235 ThrIleLeuHisPheGlnAlaLeuMetAsnIleProValIleValLeuAs 251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
618 TGCCGATCTGAACCTGGAAACATTTGCGACGATACAGCGCTGGAG 666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
251 pValAsn.....AspAspPheSerGluGluValThrIleGlnIle 264

```

```

RX MEDLINE-90324937; PubMed-2165135;
RA Tartaglia J., Winslow J., Goebel S., Johnson G.P., Taylor J.,
RA Paoletti E.;
RT "Nucleotide sequence analysis of a 10.5 kbp HindIII fragment of
RT fowlpox virus: relatedness to the central portion of the vaccinia
RT virus HindIII D region."
RL J. Gen. Virol. 71:1517-1524(1990).
CC - SIMILARITY: BELONGS TO THE DCK/DCK FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X17202; CAA35067.1; -.
DR PIR: D35216; D35216.
DR InterPro: IPR002624; -.
DR Pfam: PF01712; dmk; 1.
KW Hypothetical protein; Transferase; Kinase; ATP-binding.
FT NP_BIND 16 23 ATP (POTENTIAL).
SQ SEQUENCE 219 AA; 25908 MW; 4C3B08D85B58CE7 CRC64;

```

```

alignment_scores:
  Quality: 244.50      Length: 219
  Ratio: 1.940         Gaps: 10
  Percent Similarity: 57.534   Percent Identity: 31.507

```

alignment block:

US-09-416-579a-1 x YF26_FOWP1 ..

Align seg 1/1 to: YF26_FOWP1 from: 1 to: 219

```

73 ATCGAGGCGAACATCGGACGCGGAGACCGACGATTTGACACACTTGA 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
14 lIeGluGlyAsnIleSerSerGlyIleThrIleAspValLeuAsnIleLeuArg 30
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
123 GAAGTACACAGAAC.....GACATTTGCGCTGACCGAGC 157
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30 gAsnIleAsnAsnValValSerPheHisAspValGluAspArgTyrThrP 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
158 CGCTGCGAGGTGGCGCAACGTCACAGGGGTAAATGCTGTGAGCTGATG 207
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
47 rIleGluGlyAsn.....GluLeuIleArgIlePhe 56
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
208 TACAAAGATCCCAAGAGTGGGCGCATGCCCTTTCAGAGTATGTCACGCT 257
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
57 HisGluAsnProSerArgTyrPheSerTyrAlaLeuGlnThrHisTyrCysMe 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
258 GACCATGCTGCGAGTGCACACC.....GCCCCAACCAAGAAGAAC 298
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
73 tIysArgValArgMetHisLeuGluCysPheValProSer.....ArgV 88
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
299 TAAAAATTAAGGAGCGCTCCATTTTACGCGCTGCGTATGCTGCTGAG 348
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
88 aLAsnIleLeuGluArgSerIlePheSerAspArgTyrValPheAlaGlu 104
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
349 AACATCGCAGCAAGCGCTGCTG.....GACGAGGCGATGTACAA 389
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
105 AlaAlaThrAlaLeuGlyTyrMetAspAspProGluIleThrAlaLeuTyrCy 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
390 TACGCTGAGGAGGTGTATCAACTTCATGCAAGATCCATTACCTGCGAGG 439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 sIysGlnHisAspTyrTyr.....ThrAspIleLeuGlnIleGlnP 135
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
440 CGGACCTCATATATATGCGACCTGCGGAGGAGTGGCGAGGACGACG 489
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
135 heAspGlyIleIleTyrIleuArgThrIleProGluIleCysIleGluArg 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
490 ATCGGCGACGCGGCTGCTTTCAGAGAGAGTGC.....GTGCGCGCTTAA 533

```


152	ILASmCtIuLysSerIleThrgIuLysAsnTyrProAsnIleSerIleAs	168
534	GTACCTTTGAGAGCTGCATGATTCACGACAGAGCGTGTGATTAACACAGA	583
168	PTyLeuLysTyrIleuHisGluLysHisGluLeuTyrLeu.....	181
584	GACGACCGCAGTGTGCAGG.....GTGCTAGTCTCTGATGCCGAT	624
182ThrgInCysLysLysValProValIleuIleIleAspGlu	195
625	CTGAACTGTGAAACATTGTGCACCGCATACACAGCGCTGGAGACACAGAT	674
196GluAspPheIle	199
675	ATTTCGAC	681
199	ePheAsp	201

```

seq_documentation_block:
ID KITH_HSV1 STANDARD: PRT: 228 AA.
AC P28855:
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE THYMIDINE KINASE (EC 2.7.1.21).
GN TK OR 5.
OS Ictaluriid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae.
OC unclassified Herpesviridae.
OX NCBI_taxid=10401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AU8URN 1;
RC MEDLINE=92087490; PubMed=1727613;
RA Davison A.J.;
RT "Channel catfish virus: a new type of herpesvirus.";
RL Virology 186:9-14(1992).
RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=92013982; PubMed=1919533;
RA Harrison P.T., Thompson R., Davison A.J.;
RT "Evolution of herpesvirus thymidine kinases from cellular
  deoxycytidine kinase.";
RL J. Gen. Virol. 72:2583-2586(1991).
CC -1- CATALYTIC ACTIVITY: ATP + THYMIDINE = ADP + THYMIDINE
  5-PHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE DCK/DCK FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M75136; AAA88186.1; -.
DR EMBL: M75136; AAA88108.1; -.
DR PIR: J01336; KIBEIC.
DR InterPro: IPR002624; -.
DR Pfam: PF01712; GNM; 1.
KW Transferase; kinase; DNA synthesis; ATP-binding.
FT NP_BIND 23 30 ATP (PROBABLY).
SQ SEQUENCE 228 AA: 25642 MW: 980C26879787280B CRC64.

```

```

alignment_block:
US-09-416-579A-1 x KITH_HSVI1 ..
Align seg 1/1 to: KITH_HSVI1 from: 1 to: 228

```

46 GGGGGACCCAGCCCTTCACCGCTCTC.....AT 74
5 GUGUGYrProthrrProAlaArgPheLeuProGUGUGLeuValPheGysVa 21
75 CGAGGGCAACATCGGCAAGCGGGAAGCAACGATTATTGAACCACTTCCAGA 124
21 GGUUUGYsnIleGlyCysGlyLysSerThrLeuValLysAlaLeuMetG 38
125 AG.....TACAGAGCAACATTTGGCTGCTGACGAGCGCGGTGAGAG 168
38 LUArgValAlaGlySerGlyValAsnValAlaGUGUGProValAspG 54
169 TGGCGCAACGTCAAGGGGTAATAATTGCTGCAAGCTGATGTAAGAATCC 218
55 TPrValAsnHisAsnGlyLysAsnTyrLeuGUGLeuSerTyrThrAspPr 71
219 CAAGAAAGGGGCATGGCCCTTTCAG.....AGTTATGCA 253
71 oThrGlyTyrAlaValrProPheGlnsLeuValrPheAspSerTyrVala 88
254 CGCTGACCATGCTGACGTGCGACACGCCGCCCAACACAGAGCTAAAA 303
88 snValGluArgLeuGlnAsnPro.....Asp 96
304 ATAAAGGAGCGGCTCATTTTATGCGGTCGC.....TATTGCTCGTGA 347
97 ILeMetGluArgSerProMetSerAlaThrArgValrPheCysAlaValAs 113
348 GAACATGCGGACGA.....AACGGCTCGCTGAGACGAGGCAATGACAATA 391
113 nGlySerArgGlyValrIleProAlaThrAlaLeuProGlyMetAlaAla 130
392 CGCTGAGGAGTGGTACACAGTCAAGAGAGATGCACTTCACTGCGAGCG 441
130 rGlyGluAlaValaMetArgThrIle.....AlaThrArg 141
442 GACATCATCATATATCTGGCCACCTCGCGGAGAGTGGCGGTACAGACGAT 491
142 ProValrPheValrTyrLeuGlnLeuProProGUGUGCysLeuArgGMe 158
492 CGGGAGAGGGGCGCTGCTGAGGAGAGCTGCGCGCGCTAGTACCTGCT 541
158 LArgTrArgArgAspArgThrGlyGluAlaGlyAlaGlyLeuAspTyrLeuA 175
542 AGGACGTCATGATGTGACACGAGACGATGTTG 573
:: ||||||| ::|||
175 rGleuLeuHisGluArgTyrGluAlaTrrPLeu 185
seq_name: SwissProt_39:KITH_HSVSA
seq_documentation_block:
ID KITH_HSVSA STANDARD: PRT: 527 AA.
AC P21293;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE THYMIDINE KINASE (EC 2.7.1.21).
GN 21 OR TK.
OS Herpesvirus saimiri (strain 11).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=10383;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90063548; PubMed=2555434;
RA Honess R.W., Craxton M.A., Williams L., Gompels U.A.;
RT "A comparative analysis of the sequence of the thymidine kinase gene"

RT of a gammaherpesvirus, herpesvirus saimiri.";
 RL J. Gen. Virol. 70:3003-3013(1989).
 [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=92333688; PubMed=1321287;
 RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Bisinger B.,
 RA Newman C., Wilmann S., Craxton M.A., Coleman H., Fleckenstein B.,
 RA Honess R.W.;
 RT "Primary structure of the herpesvirus saimiri genome.";
 RL J. Virol. 66:5047-5058(1992).
 CC -1- CATALYTIC ACTIVITY: ATP + THYMIDINE = ADP + THYMIDINE
 CC 5'-PHOSPHATE.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES THYMIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X64346; CAA45643.1;
 DR EMBL: D00543; BAA00432.1;
 DR PIR: A33374; KIBBS.
 DR InterPro: IPR001889;
 DR Pfam: PF00693; TK_herpes; 1.
 KW Transferase: Kinase; DNA synthesis: ATP-binding.
 FT NP_BIND 216 223 ATP (BY SIMILARITY).
 SEQUENCE 527 AA: 59806 MW; 4EBE3A3A178FAD7D CAC64;

alignment_scores:
 Quality: 134.50 Length: 257
 Ratio: 1.027 Gaps: 13
 Percent Similarity: 50.973 Percent Identity: 25.292

alignment_block:
 US-09-416-579A-1 x KITH_HSVSA ..

Align seg 1/1 to: KITH_HSVSA from: 1 to: 527

```

37 AASTAGCCGAGGACCCACCCCTTACCGCTTATGAGGCGACACT 86
||||| : : : : : ||| : : : : : ||| : : : : : |||
202 LSHISGLMSAPROVALPROALAPHELIAPHELEUCLNLYSERILL 218
||||| : : : : : ||| : : : : : ||| : : : : : |||
87 CGGACGCGGAGACGACGATTTG.....AACCACTTGAGACAACT 127
||||| : : : : : ||| : : : : : ||| : : : : : |||
218 eslyvalgllylthrlthleuleulyssermetasnnglylleuaglyg 235
||||| : : : : : ||| : : : : : ||| : : : : : |||
128 ACAAGAACGACATTTGCTGCTGACGACCGCTGAGAGAGTGGCGCAAC 177
||||| : : : : : ||| : : : : : ||| : : : : : |||
235 lylsasn...ValleuAlaphehISgluProllealatyrtPrhAsp 250
||||| : : : : : ||| : : : : : ||| : : : : : |||
178 GTCAGCGGGTAATCTGCTGGAGCTGAGTACAA..... 213
||||| : : : : : ||| : : : : : ||| : : : : : |||
251 ValPheSer...AsnserleuGluGluValTylyLysleuThlProal 266
||||| : : : : : ||| : : : : : ||| : : : : : |||
214 .....GATCCCAAGA 223
266 alysvAlglaryGThrsAsnserAlalysleuAlaIacysGlnleu 283
||||| : : : : : ||| : : : : : ||| : : : : : |||
224 AGTGGCCATGCCCTTTCAGAGTTATGACGCTGACATGCTGACGTG 273
||||| : : : : : ||| : : : : : ||| : : : : : |||
283 yspheIaserProleuAlaIeulysThralaThr.....Aspatg 297
||||| : : : : : ||| : : : : : ||| : : : : : |||
274 CACACGCCCCCAACCAAGAAGCTAAAA.....ATAAR 308
||||| : : : : : ||| : : : : : ||| : : : : : |||
298 LeuSerSerProlysnSerleuSerSerAspmetlrrPvalMetph 314
||||| : : : : : ||| : : : : : ||| : : : : : |||
309 GGAGCGCTCATTTTTCAGCGCTGCTGCTGAGAGACATGCGCAG 358
||||| : : : : : ||| : : : : : ||| : : : : : |||
314 esapArgHISProleuSerAlaThrValValPheProTyTmetHISphg 331

```

```

359 GAAACGCTCGCTGAGCAGGACGATGATACATACCTGAGGATGCTAC 408
||||| : : : : : ||| : : : : : ||| : : : : : |||
331 lnaSnnglyPheleu.....SerPheSerHISleuIleGlnleuTrpSer 345
||||| : : : : : ||| : : : : : ||| : : : : : |||
409 AAGTTCATGCAAGAGCTCATTCACCTGACGCGGACCTCATATATCT 458
||||| : : : : : ||| : : : : : ||| : : : : : |||
346 SerPheIysAlaser.....ArgGlyAspAsnIleIleleu 358
||||| : : : : : ||| : : : : : ||| : : : : : |||
459 GGGACACCTCGCCGAGAGTGGCTACGACGATCCGCGAGCGGCTGCT 508
||||| : : : : : ||| : : : : : ||| : : : : : |||
358 uasnleuasnserGlnGluasnleuIysArgValIysIysArgasnArgL 375
||||| : : : : : ||| : : : : : ||| : : : : : |||
509 CTGAGAGAGAGCTGCTGCCGCTTACCTTACCTTACGAGAGCTG..... 549
||||| : : : : : ||| : : : : : ||| : : : : : |||
375 ysgluGlnIySerAlaserIleGlnHISleuAlaGlyleuAsnAsnGly 391
||||| : : : : : ||| : : : : : ||| : : : : : |||
550 ..CATGAGTTGCACACGAGACTGTTGATACACACAGACA...CGACCGCA 593
||||| : : : : : ||| : : : : : ||| : : : : : |||
392 TyrlHISAlaValTylyCysAlaTyrlProleuValGlnAsnPhetlProgl 408
||||| : : : : : ||| : : : : : ||| : : : : : |||
594 GTCGTGCAAGTCTTACCTTCGATGCC.....GATCTGAC. 630
||||| : : : : : ||| : : : : : ||| : : : : : |||
408 uGluIleValGluValIyCysPheasnAlaIySHISleThrasPheusers 425
||||| : : : : : ||| : : : : : ||| : : : : : |||
631 .....CTGMAAACATTGGCAGCAGCAGTACGACGCG 660
||||| : : : : : ||| : : : : : ||| : : : : : |||
425 erSerlyProSerPheleuAlaIySHISValSerThrlGlu...Aspmet 440
||||| : : : : : ||| : : : : : ||| : : : : : |||
661 TCGGAGACGACGATTCGAC 681
||||| : : : : : ||| : : : : : ||| : : : : : |||
441 leuIysSerSerIlePheasn 447
||||| : : : : : ||| : : : : : ||| : : : : : |||

```

seq_name: Swissprot_39:DGK2_LACAC

seq_documentation_block:
 ID DGK2_LACAC STANDARD; PRT; 223 AA.

```

AC 059484;
DT 15-JUL-1998 (Rel. 36, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DEOXYGUANOSINE KINASE (EC 2.7.1.113) (DGUO KINASE) (DGK)
DE (DEOXYNUCLEOSIDE KINASE COMPLEX I F-COMPONENT).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1579;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11506 / R-26;
RX MEDLINE=95204449; PubMed=7896798;
RA Ma G.-T., Hong Y.S., Ives D.H.;
RT "Cloning and expression of the heterodimeric deoxyguanosine
  kinase/deoxyadenosine kinase of Lactobacillus acidophilus R-26.";
RL J. Biol. Chem. 270:6595-6601(1995).
RN [2]
RP SEQUENCE OF 1-17.
RC STRAIN=ATCC 11506 / R-26;
RX MEDLINE=94227067; PubMed=8172906;
RA Ikeda S., Ma G.-T., Ives D.H.;
RT "Heterodimeric deoxynucleoside kinases of Lactobacillus acidophilus
  R-26: functional assignment of subunits using limited proteolysis
  controlled by end-product inhibitors.";
RL Biochemistry 33:5328-5334(1994).
CC -1- FUNCTION: DGK/DAK PLAYS AN ESSENTIAL ROLE IN GENERATING THE
  DEOXYRIBONUCLEOTIDE PRECURSORS, DCTP AND DATP, FOR DNA METABOLISM.
CC -1- CATALYTIC ACTIVITY: ATP + DEOXYGUANOSINE = ADP + GMP.
CC -1- SUBUNIT: HETERODIMER OF A DEOXYADENOSINE (DAK) AND A
  DEOXYGUANOSINE KINASE (DGK).
CC -1- SIMILARITY: BELONGS TO THE DGK/DGK FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U01881; AAB09751.1; -
DR InterPro; IPR002624; -
DR Pfam; PF01712; DNM; 1.
KW Transferase; Kinase; ATP-binding.
FT INIT MET 0
FT NP_BIND 7 14 ATP (POTENTIAL).
SQ SEQUENCE 223 AA; 26183 MW; 87EF74968751ED5D CRC64;

alignment_scores:

Quality: 134.00 Length: 212
Ratio: 1.039 Gaps: 8
Percent Similarity: 60.849 Percent Identity: 22.642

alignment_block:

US-09-416-579A-1 x DGR2_LACAC

Align seg 1/1 to: DGR2_LACAC from: 1 to: 223

```

67 GTCTCATCGAGGGCGACATCGCGAGCGGAGACCACTATTGACCA 116
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   3 lIeValleuSerGlyProIleGlyAlaGlySerSerleuThrGlyI 19
117 CTTCGAGAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 166
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   19 eIeuSerIeTyTyr.....LeuGlyThraSnpTrohe.... 29
167 AGTGGCCACAGCTACAGCGGGTAAATCTGCTGAGCTGATGACAA 216
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   30 ..TyGluSerValaSpaPasnProValleuProleuPheTyGluAsn 45
217 CCCAAGAGTGGCGCATGCCCTTTCAAGATTATGTCACGCTGACCA 266
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   46 ProIySlyTyAlaPheLeuLeuGluValaTyPheLeuasnThrargPh 62
267 GCACTGCCACACCGCCCAACACAGAGAGAGAGAGAGAGAGAG 316
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   62 earGSeTlleYSerIleAlaLeuThrAspaPasnValleuSparG 79
317 CCATTTTACGCGCTGCTATTGCTCTGCGAGAGAGAGAGAGAG 366
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   79 eIleTyGluAspAlaLeuPhePhe.....GlnMetaSnaIa 91
367 TCCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 416
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   92 AsplleGlyTrgAlaThrProGluGluValaSpThrTyTrgIleuLe 108
417 CGAAGAGTCCATT.....CACCTGACAGCGGAGC 445
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
108 uHlSaSmeMetSerGluLeuasnParGmetProIySlySaSnpProSPL 125
446 TCATCATATATCTGCGACCTCGCGAGAGTGGCGTACGAACCAT 495
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
125 lIeueuValaHlSIlleAspValSerTyAspThrMetLeuIySargIleGln 141
496 CAGCGGCGCTGCTGTGAGAGAGAGTGCCTCCCTTAAGACTTGA 545
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
142 LySargIlyTrgAsnTyTrgIleuGln.....LeuSerTyTrgAspProTh 155
546 GGTGCATGAGTTGACACGAGACTGTTGATACACAGAGAGAGAG 595
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
155 rIeugIuAspTyTrgTyTrgIySarg...LeuEuargTyTrgIySproTript 171
596 CGTGCAGAGTC.....CTAGTCCATGATCCGATCTGAAC 630
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
171 yTrAlaIySlyTyAspTySerProIySmetThrIleAspGlyAsp...Lys 186

```

631 CTGGAACAATTGGCACCGGAGTACACCGCTCGGAG 666
187 LeuaspPheMetAlaSerGluGluAspArgGln 198

seq_name: SwissProt_39:YAAF_BACSU

seq_documentation_block:

ID YAAF_BACSU STANDARD; PRT; 217 AA.
AC P37529;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE HYPOTHETICAL 25.4 KDA PROTEIN IN SENS-DNAZ INTERGENIC REGION.
GN YAAF.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxId=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96051385; PubMed=7584024;
RA Ogatawara N., Nakai S., Yoshikawa H.;
RT *Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin.";
RL DNA Res. 1:1-14(1994).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D26185; BAA05250.1; -
DR EMBL; Z99104; CAB11790.1; -
DR Subtilist; BG10078; YAAF.
DR InterPro; IPR002624; -
DR Pfam; PF01712; DNM; 1.
KW Hypothetical protein.
SQ SEQUENCE 217 AA; 25444 MW; B02CDBAA1C518305 CRC64;

alignment_scores:

Quality: 128.50 Length: 223
Ratio: 1.117 Gaps: 7
Percent Similarity: 51.570 Percent Identity: 23.318

alignment_block:

US-09-416-579A-1 x YAAF_BACSU

Align seg 1/1 to: YAAF_BACSU from: 1 to: 217

```

67 GTCTCATCGAGGGCGACATCGCGAGCGGAGAGAGAGAGAGAG 116
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   12 lIeThValaIaGlyThrValGlyAlaGlySerThrleuThrGlyTh 28
117 CTTCGAGAG.....TACAGAGAGAGAGAGAGAGAGAGAGAG 157
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   28 rIeuAlaIySargIleuGlyPheTyTrgSerLeu..... 39
158 CGTGCAGAGTGGCGGACAGCTACAGCGGGGTAAATCTGCTGAGCTG 207
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   40 .....GluGluValaIaSpHlSaSnpProTyTrgLeuGluIySfhe 51
208 TACAGAGATCCCAAGAGTGGCGCATGCCCTTTCAGAGATTATGCA 257
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   52 TyHlSaSnpPheGluArgTySerPheHlSleuGlnIleTy..... 65
258 GACCATGCTGCGAGTGGCGACACCGCCCAACAGAGAGAGTAA 303
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   66 .....PheleuAlaGluArgPheIySleuGlnIySfThrI 77

```


seq_name: SwlssProt_39:YHL1_EBV

RP SEQUENCE FROM N.A.

RX MEDLINE-84270667; PubMed-6087149;
 RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
 RA Gibson T.J., Hatfull G., Hudson G.S., Satchell S.C., Seguin C.,
 RA Tuftnell P.S., Barrett B.G.;
 RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome." *Nature* 310:207-211(1984).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR	EMBL	PIR	AA	NOT_ANNOTATED	CDS
DR	U01555	-	-	NOT_ANNOTATED	CDS
DR	A03742	Q0853	149	1	1
KM	Hypothetical protein; Early protein; Repeat.		4 X 125	AA	TANDEM REPEATS
ET	DOMAIN	149	648	1	1
ET	REPEAT	149	273	1	1
ET	REPEAT	274	398	2	2
ET	REPEAT	399	523	3	3
ET	REPEAT	524	648	4	4
SEQUENCE	650 AA;	86DA1D67A37152A2	CRC64		

[illegible]

alignment_block:
US-09-416-579A-1 x YHLL_EBV

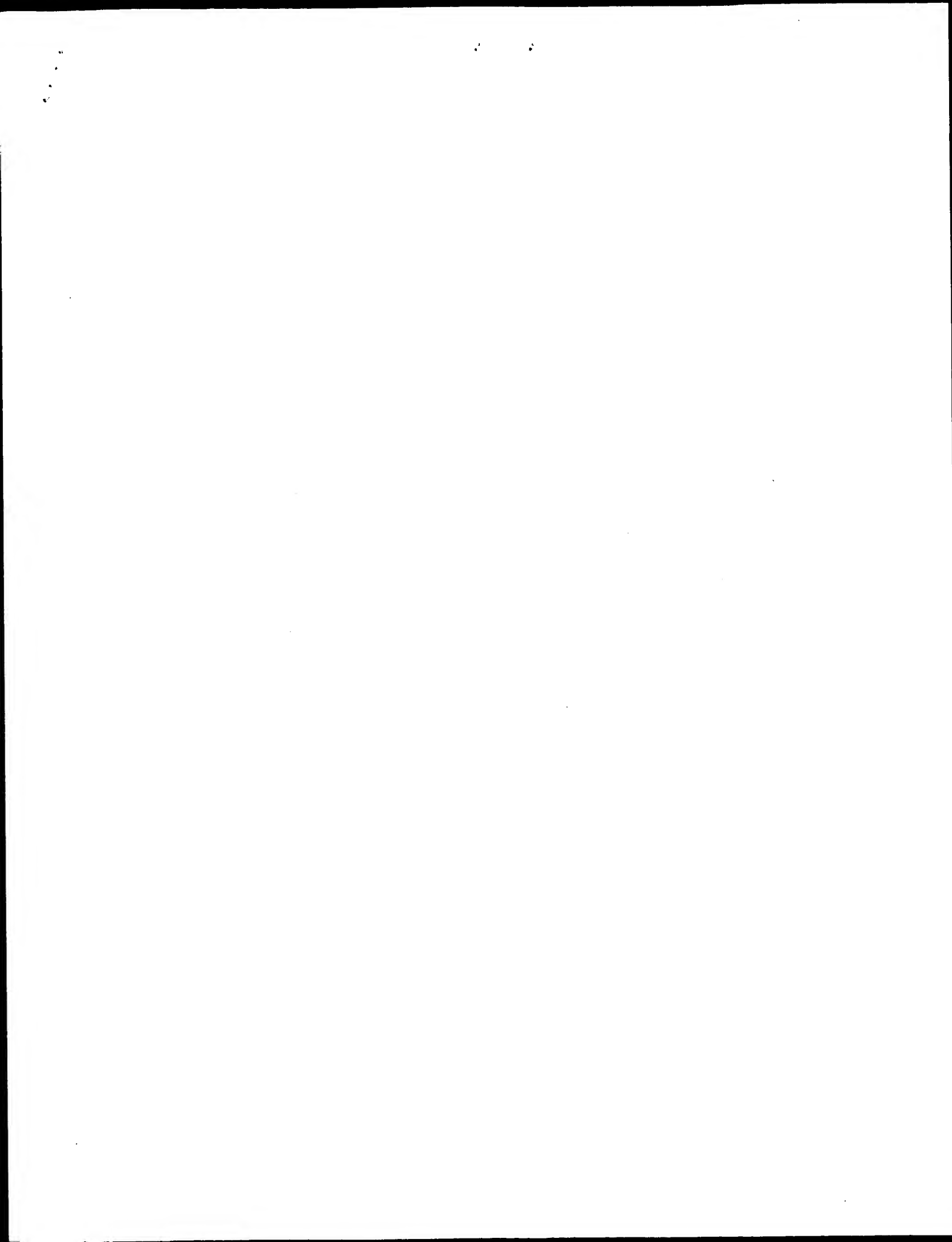
Align seg 1/1 to: YHL1_EBV from: 1 to: 660

```

2  TGGCGGAGGACAG.....CATCTGTGGCCCGAAGGG 33
   |||||
159  TGTATGATGATGSeCtYlaGlaGlnaGtYlnaSProPfoGtYlaGcl 175
34  A.....CCAAATAGCGGACGAGGACCCAGGCGCTTACAGCGTCATGCA 77
   |||||
175  yGlnaGtPfoSeRgYlPfoThhGtYlaGtPfoGlaMaLaPfoGtYlaP 192

```

[illegible]




```

158  TPhHisAspTrpMetAsnSerGlnPheGlyGlnSerLeuGlnLeuAspG1 174
447  CATCATATATCTGGACACCTGGCCGAGGTGGCTAGCAAGCATCCGGC 436
      |||||||
174  YLlelLeuLeuAlaGlnAlaThrProGlnLysCysLeuAsnArg1LeuYL 191
497  AGCGGCTCTTCTGAGAGAGAGCGGCTGAGTACCTTCAGAGAG 546
      |||||||
191  euArgGlyArgAsnGlnGlnGlnGlyLeuProLeuGlnLysLeuGln 207
547  CTGCATGAGTTGCACAGACTGGTGTGATACACAGAGACAGCGGAGTC 596
      |||||||
208  LeuHisTrpTrpLysHisGlnSerTrpLeuLeuHisArgThrLeuLys 224
597  GTGC.....AAGTCTGATCTCGATGCCGATCTGAC 630
      |||||||
224  rPhAspTrpLeuGlnGlnLysValProValLeuThrLeuAspValAsn 239
seq_name: p1r1.A38585

```

```

seq_documentation_block:
deoxycytidine kinase (EC 2.7.1.74) - human
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A38585; S14321
R:Chottiner, E.G.; Shewach, D.S.; Datta, N.S.; Ashcraft, E.; Gibblin, D.; Ginsburg, D.;
Proc. Natl. Acad. Sci. U.S.A. 88, 1531-1535, 1991
A:Title: Cloning and expression of human deoxycytidine kinase cDNA.
A:Reference number: A38585; M0ID:91142207
A:Accession: A38585
A:Molecule type: mRNA
A:Residues: 1-260 <CHO>
A:Cross-references: GB:M60527; NID:9181509; PIDN:AAA35752.1; PID:9181510
R:Eriksson, S.; Cederlund, E.; Bergman, T.; Joernvall, H.; Bohman, C.
FEBS Lett. 280, 363-366, 1991
A:Title: Characterization of human deoxycytidine kinase. Correlation with cDNA sequences
A:Reference number: S14321; M0ID:91192170
A:Accession: S14321
A:Molecule type: protein
A:Residues: 35-44, 'X', 46-47, 'X', 49-56; 163-168, 'X', 170-181; 223-236 <ERT>
A>Note: the amino end is blocked
C:Genetics:
A:Gene: GDB:DCX
A:Cross-references: GDB:126810; OMIM:125450
A:Map position: 4q13.3-4q21.1
A:Superfamily: human deoxycytidine kinase
C:Keywords: ATP; blocked amino end; homodimer; phosphotransferase

```

```

alignment_scores:
Quality: 356.50 Length: 247
Ratio: 2.214 Gaps: 5
Percent Similarity: 65.182 Percent Identity: 30.769

```

alignment_block:

US-09-416-579A-1 x A38585 ..

Align seg 1/1 to: A38585 from: 1 to: 260

```

43  GCCGAGGACCCACCCCTTCACGCTCTCATCGAGGGCAACATCCGAG 92
      |||||||
16  SerGlnGlyThrArg1LeuLysLysIleSer1IleGlnLysAsn1LeuAla1 32
93  CGGGAAGACACGATTTGACCACTTCGAGAGATACAAAGACGACATT 142
      |||||||
32  ad1LysSerThrPheValAsn1IleLeuLysGlnLeuLysGlnAspTrpG 49
143  GCCTGCTGACCGAGCCGCTCGAGAAAGTGGCGCAACGTC..... 180
      |||||||
49  LuValValProGlnProValAlaArgTrpCysAsnValGlnSerThrGln 65
181  .....AACGGGCTAAATCTGCT 197
      |||||||

```

```

66  AspGlnPheGlnGlnLeuThrMetSerGlnLysAsnGlnLysValLe 82
198  GGACGTGATGTACAAGATCCCAAGAGAGTGGGCAATGCCCTTGAGAGTT 247
      |||||||
82  uGlnMetMetLysGlnLysProGlnArgTrpSerPheThrPheGlnThr 99
248  ATGTACAGCTGACACATGCTGCGAGTGCACACGCCGCCCAACAGCAAG 297
      |||||||
99  yAlaLysLeuSerArg1LeuArg1AlaGlnLeuAlaSerLeuAsnGln 115
298  CTAAAA.....ATTAAGGAGCGCTCCATTTTAG 326
      |||||||
116  LeuLysAspAlaGlnLysProValLeuPhePheGlnValArgSerVal 132
327  CGCTGCTATGCTTGTGAGAGACATGCGAGAAACGGCTGCGTGGAGC 376
      |||||||
132  rAspArgTrpTrpLeuPheAlaSerAsnLeuTrpGlnSerGlnCysMet 149
377  AGGCGATGTACAATACGCTGAGAGAGTGTACAAAGTTCATCGAAGATCC 426
      |||||||
149  LuThrGlnTrpThrLeuTrpGlnAspTrpHisAspTrpMetAsnGln 165
427  ATT.....CACTGACGCGGACCTCATCATATATCTCCGACCTCGCC 470
      |||||||
166  PheGlnLysSerLeuGlnLysAspGlnLysLeuLeuGlnAlaThrPr 182
471  GGAGTGGGCGGTACAGACGCGTCCGAGCGGCTGCTTGAGAGAGCT 520
      |||||||
182  oGlnTrpCysLeuHisArg1IleTrpLeuArgGlnArgAsnGlnGlnG 199
521  GCGTGGCGCTTAAGTACCTTCACAGAGCTGCATGAGTTGCACAGAGTCG 570
      |||||||
199  LysLeuProLeuGlnLysValGlnLysLeuHisTrpLysHisGlnSer 215
571  TTGATCACACGAGAGACGCG.....CACTGCGCAA 602
      |||||||
216  LeuLeuHisArgThrLeuLysThrAsnPheAspTrpLeuGlnGlnVal 232
603  GGTCTAGTCTGATGCGGATCGGATCTGAACCTGGAAACATTTGGCACCAGT 652
      |||||||
232  oLleuLeuThrLeuAspValAsn.....GluAspPheLysAspLysT 246
653  ACCAGCGCTCGGAGACGACGATTTGACGCGCATCTCAAGT 693
      |||||||
246  yrgLysSerLeuValGlnLysValLysGlnPheLeuSerThr 259
seq_name: p1r2.JC6142

```

seq_documentation_block:

```

deoxyguanosine kinase (EC 2.7.1.113) precursor - human
C:Species: Homo sapiens (man)
C>Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 18-Jun-1999
C:Accession: J06142
R:Johansson, M.; Karlsson, A.
Proc. Natl. Acad. Sci. U.S.A. 93, 7258-7262, 1996
A:Title: Cloning and expression of human deoxyguanosine kinase cDNA.
A:Reference number: J06142; M0ID:96293511
A:Accession: J06142
A:Molecule type: mRNA
A:Residues: 1-277 <JOH>
A:Cross-references: GB:U15668; NID:91477481; PIDN:AAC50624.1; PID:91477482
A>Note: It is uncertain whether Met-1 or Met-8 is the initiator
C:Comment: This enzyme is involved in mediating cytotoxicity of nucleoside analogs. I
      inase, thymidine kinase 2.
C:Superfamily: human deoxycytidine kinase
C:Keywords: phosphotransferase
F:1-39/Domains: signal sequence #status predicted <SIG>
F:40-277/Product: deoxyguanosine kinase #status predicted <MAT>

```

```

alignment_scores:
Quality: 340.00 Length: 233
Ratio: 2.329 Gaps: 5

```


Percent Similarity: 62.661 Percent Identity: 33.047

alignment_block:

US-09-416-579A-1 x JG6142 ..

Align seg 1/1 to: JG6142 from: 1 to: 277

```

49 GGCACCCAGCCCTTCACCGCTTCATCGAGGCGCAACATCCGCGACGGGAA 98
   ||| ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
35 G1YArgG1ProArgArgLeuSerIleGluGlyAsnIleAlaValGlyLy 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
99 GACCAAGATTTGAACCTTCGAGATCAAGAACAGCAATTCCTGC 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 SsetThrPheValIysLeuLeuThrIlystIlyrProGluThrPheIstVala 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
149 TGACCGAGCCCGTCGAGAAAGTGGCGCAACGTCAAGGGGTA..... 189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 IatHrGluProValAlaThrTrpGlnAsnIleGlnAlaAlaGlyAsnGln 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
190 .....AATCTGCTGAGCGAGTCAAA 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
85 LysAlaCysThrAlaGlnSerLeuGlyAsnLeuLeuAspMetCtyrAr 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
213 AGATCCCAAGAGTGGCCATGCCCTTCAGAGTATATGACGCTGACCA 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
101 GGluproAlaArgTrpSerIlyrThrPheGlnThrPheSerPheLeuSera 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
263 TGCTGCAGCTGCACACCGCCCAACCAAGAAAGCTA..... 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
118 rGleuValGlnLeuGluProPheProGluIlySLeuLeuGlnAlaArg 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 .....AATTAARGAGCGCTCCATTTTTCAGCGCTGCTATGCTT 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
135 LysProValGlnIlePheGlnArgSerValIlySerAspArgIlyrIlePh 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
342 CGTGGAGAACATGCGGACGAGGAGCGCTGAGGAGCGGAGTGTACATA 391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
151 eAlaLysAsnLeuPheGlnAsnGlySerLeuSerAspIleGluThrPheIst 168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
392 CGCTGGAGAGTGTACAGTTTCATC.....GAAGAGTCCATTCACCTG 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
168 IeTyGlnAspTrpHisSerPheLeuThrTrpGluPheAlaSerArgIle 184
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
436 CAGCGGAGCTCATCATATATCTGCGACCTCGCGGAGGTGGCGTACGA 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
185 ThrLeuHisGlyPheIleTyLeuGlnAlaSerProGlnValCysLeuLy 201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
486 ACCGATCCGAGCGGCGCTGCTGAGAGAGAGTGCCTGCCGCTTAAGT 535
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
201 sArgLeuTyGlnArgAlaArgGluGluGlyIleGluLeuAlaIAT 218
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
536 ACCTTCAGAGAGTGCATGATGTCACCGAGACTGTTGATACACCAAGGA 585
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
218 rYLeuGluGlnLeuHisGlyGlnHisGluAlaTrpLeuIleHisIlyThr 234
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
586 CGACCG.....CAGTCGTGCAAGTCTCTAGTCTTACGA 617
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
235 ThrIlySLeuHisPheGluAlaLeuMetAsnIleProValIleuValLeuAs 251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
618 TGCCGATGTGAACCTGGAAGAACATTCGACCGAGTACCGAGCGCTGAG 666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
251 rValAsn.....AspAspPheSerGluValThrIlySglnGln 264
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
seq_name: p1r2:S71315
seq_documentation_block:
deoxyguanosine kinase (EC 2.7.1.113) precursor - human
C:Species: Homo sapiens (man)
C:Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #extl_change 20-Jun-2000
C:Accession: S71315; S78432
R:Mang, L.; Hellman, U.; Eriksson, S.
FEBS Lett. 390, 39-43, 1996
A:Title: Cloning and expression of human mitochondrial deoxyguanosine kinase cDNA.

```

A:Reference number: S71315; MUID:96314545

A:Accession: S71315

A:Molecule type: mRNA

A:Residues: 1-260 <MAN>

A:Cross-references: EMBL:X97386; PIDN:CAA66054.1

A:Experimental source: tissue brain

R:Mang, L.; Hellman, U.; Eriksson, S.

submitted to the EMBL Data Library, April 1996

A:Description: Cloning and expression of human deoxyguanosine kinase cDNA.

A:Reference number: S78432

A:Accession: S78432

A:Molecule type: mRNA

A:Residues: 1-18,'R',20-260 <MAN>

A:Cross-references: EMBL:X97386; PIDN:CAA66054.1

A:Experimental source: brain

C:Genetics:

A:Gene: dgk

C:Superfamily: nuclear

C:Key words: human deoxycytidine kinase

F:1-32/Domain: transit peptide (mitochondrion) #status predicted <TMP>

F:23-260/Product: deoxyguanosine kinase #status predicted <MAT>

F:125-132/Region: DKS motif

F:185-191/Region: arginine-rich

alignment_scores:

Quality: 333.00 Length: 233
Ratio: 2.297 Gaps: 5
Percent Similarity: 62.232 Percent Identity: 32.618

alignment_block:

US-09-416-579A-1 x S71315 ..

Align seg 1/1 to: S71315 from: 1 to: 260

```

49 GGCACCCAGCCCTTCACCGCTTCATCGAGGCGCAACATCCGCGACGGGAA 98
   ||| ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
18 G1YArgG1ProArgArgLeuSerIleGluGlyAsnIleAlaValGlyLy 34
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
99 GACCAAGATTTGAACCTTCGAGATCAAGAACAGCAATTCCTGC 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
34 SsetThrPheValIysLeuLeuThrIlystIlyrProGluThrPheIstVala 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
149 TGACCGAGCCCGTCGAGAAAGTGGCGCAACGTCAAGGGGTA..... 189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 IatHrGluProValAlaThrTrpGlnAsnIleGlnAlaAlaGlyThrGln 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
190 .....AATCTGCTGAGCGAGTCAAA 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 LysAlaCysThrAlaGlnSerLeuGlyAsnLeuLeuAspMetCtyrAr 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
213 AGATCCCAAGAGTGGCCATGCCCTTCAGAGTATATGACGCTGACCA 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
84 GGluproAlaArgTrpSerIlyrThrPheGlnThrPheSerPheLeuSera 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
263 TGCTGCAGCTGCACACCGCCCAACCAAGAAAGCTA..... 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
101 rGleuValGlnLeuGluProPheProGluIlySLeuLeuGlnAlaIAT 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 .....AATTAARGAGCGCTCCATTTTTCAGCGCTGCTATGCTT 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
118 LysProValGlnIlePheGlnArgSerValIlyrSerAspArgIlyrIlePh 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
342 CGTGGAGAACATGCGGACGAGGAGCGCTGAGGAGCGGAGTGTACATA 391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
134 eAlaLysAsnLeuPheGlnAsnGlySerLeuSerAspIleGluThrPheIst 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
392 CGCTGGAGAGTGTACAGTTTCATC.....GAAGAGTCCATTCACCTG 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
151 IeTyGlnAspTrpHisSerPheLeuThrTrpGluPheAlaSerArgIle 167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
436 CAGCGGAGCTCATCATATATCTGCGACCTCGCGGAGGTGGCGTACGA 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Journal of Management Inquiry 23(4) 409–424 © The Author(s) 2014. Reprints and permissions: sagepub.com/journalsPermissions.nav DOI: 10.1177/1056492614564611

probable thymidine


```

149 sTyrtyrSer.....llepheAsnlystrhrAsnSerThrProI 163
      :          :||| :          :|||
      :          :||| :          :|||
605 TCCTAGTCATGCGCCGATCGACCTGAGACCTGGAACACATGCGCCGACGATAC 654
      :          :||| :          :|||
163 lleuSerPheleAsnSerLysAsnGluAspGluValGluGluAspPhe 179
seq_name: p1r1:T42934

```

seq_documentation_block:

```

thymidine kinase (EC 2.7.1.21) [similarity] - ateline herpesvirus 3 (strain 73)
C:Species: ateline herpesvirus 3
A:Variety: strain 73
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000
C:Accession: T42934
R:Albrecht, J.C.; Fleckenstein, B.
  submitted to the EMBL Data Library, August 1998
A:Description: Primary structure of the herpesvirus ateles genome.
A:Reference number: Z22274
A:Accession: T42934
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-527 <ALB>
A:Cross-references: EMBL:AF083424; PIDN:AC95545.1
A:Experimental source: strain 73
C:Genetics:
A:Note: orf 20
C:Superfamily: salmistrine herpesvirus 1 thymidine kinase; herpesvirus thymidine kinase
C:Keywords: phosphotransferase

```

alignment_scores:

```

Quality: 149.50      Length: 254
Ratio: 1.133         Gaps: 10
Percent Similarity: 51.969      Percent Identity: 25.197

```

alignment_block:

US-09-416-579A-1 x T42934

-Align seg 1/1 to: T42934 from: 1 to: 527

```

40  TACGCGGAGCGGACCCACCTTCACCTCTCTCATCGAGGCGACATCGG 89
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
203 HisGlnGluProIleProAlaPhePheleuGluGluSerIleG1 219
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
90  CAGCGGAGACACACGATATTGAACACCTTCGAGAGATAC.....A 130
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
219 yValGlyLysThrIleuLeuLysSerMetArgGlyIleuLeuProGlyL 236
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
131 AGAAGCATTTGCTCTCTGACGAGCCCGTCGAGAGTGGCGGACATGC 180
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
236 ysaAn...ValIeuThrPheHisGluProMetAlaPheTrpLysAsnVal 251
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
181 AACGGGTAAATCTGTCGAGCTATGATACAAA.....GATCCCAAGAA 224
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
252 PheSer...AsnSerLeuAspGluValTyrLysLeuThrLeuProAla 267
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
225 GTGGGCGATGCGCTTTCAGATTATGTCACGCTGACCATGCTGAGTGC 274
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
267 sValGlySerMetThrAsnSerLysLeuLeuAlaCysGlnLeuLysP 284
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
275 ACACCGGCCCA.....ACCAACAAGAAAGCTAAATA 306
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
284 heAlaAlaProLeuLeuAlaLeuLysThrSerThrAspArgLeuSer 300
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
307 ARGAGCGCTCATTTTACGCGCTGCTATTTGCTTGGTGGAGAAATCGC 356
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
301 SerAsnArgSerAsnLeuSerSerAsnMetTrpValMetPheAspArg 317
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
357 ACGAACGCGCTCGCTGAGAGCGGACGATGACAAATACCGCTGAGAGAG 406
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
317 sProLeuSerAlaThrValValPheProTyrMetHisTyrGlnAsnGly 334

```

seq_name: p1r2:F75535

seq_documentation_block:

```

deoxyguanosine kinase/deoxyadenosine kinase subunit - Deinococcus radiodurans (strain
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: F75535
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
  S.; Smith, H.O.; Venter, J.C.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
  Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: F75535
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-207 <WHI>
A:Cross-references: GB:AE001891; GB:AE000513; NID:96457973; PIDN:AAF09882.1; PID:9645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0298
A:Map position: 1
C:Superfamily: lactobacillus acidophilus deoxyadenosine kinase

```

alignment_scores:

```

Quality: 148.00      Length: 188
Ratio: 1.345         Gaps: 6
Percent Similarity: 58.511      Percent Identity: 26.064

```

alignment_block:

US-09-416-579A-1 x F75535

-Align seg 1/1 to: F75535 from: 1 to: 207

```

67  GTCCTCATGAGGCAACATCGGACGCGAGAGACGATATTGACCA 116
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
3  ValAlaValSerGlyAsnIleGlySerGlyLysSerThrLeuThrArgMe 19
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
117 CTTTC...GAGAGTACACAGACGACATTTGCTGCTGACGAGCCGCTCG 163
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

```

19 tleuAlaGluArgTyrGly.....LeuArgProValT 30
164 AGAAGTGGCGCAAGCTCAACGGGGTAAATCTGCTGACGTATGTACAAA 213
   ::::: ||| ||| ||| ||| :::::
30 yrgLupProTyrAlaGluAsn.....ProTyrLeuGluAspPheTyrHis 44
214 GATCCCAAGAAAGTGGGCGCATGCCCTTTCAGAGTATGTACAGCTGACCAT 263
   ||| ::::: ||| ||| ||| |||
45 AspMetAlaGlnTyrSerPheHisSerGlnValTyrPheLeuSerArgAr 61
264 GCTGCAAGTGCACACCGCC...CCAACCAACAAGAGCTAAATTAARGG 310
   ||| ::::: ||| ||| ||| |||
61 gLeuGlnGlnHisLeuGlyMetValThrGlyAlaArgTyrValIleGln 78
311 AGCGCTCCATTTTAAAGCTGCTATGCTTGTGTGGAGAACATGGCAGCA 360
   ::| ||| ::| ||| ::| ||| ::| |||
78 sPArgThrValPheGluAspAlaAsnIlePheAlaArgAsnLeuTyrGlu 94
361 AACGGCTGCTGAGCAGCGCATGTACATATGACGTGAGAGAGTGGTACAA 410
   ::| ||| ::| ||| ::| ||| ::| |||
95 SerGlyIleMetGlyGlnArgAspTrpGlnThrTyrArgGlyLeuTyrCl 111
411 GTTCATCGAAGATCCATTACCTGACGGCGGACCTCATATATATCTGC 460
   : :: ::::: ||| ||| ||| ||| |||
111 uGlyValLeuProAlaLeuArgVal...ProAspLeuIleHisIleA 127
127 sPAlaGlyLeuProThrLeuArgArgArgIleAlaLeuArgGlyArgAsp 143
511 GAGGAGAGCTGCGTGGCTTAAGCTTTCAGAGCTCATAGTATGTGA 560
   ||| ::::: ||| ||| ||| ||| |||
144 TyrGlnGlnAlaIleProAspLeuTyrLeuAlaGlyLeuAsnArgLeuTyr 160
561 CCAGAGCTGGTGAATACACCAAGAGACCGAGCTGTGCAGAGTCTAG 610
   : ||| ::| ||| ||| ||| |||
160 rAlaGlyTyrPhe.....AlaAlaPheAspLeuGlyProValValA 174
611 TCCTGATGCCGAT 624
174 rGValAspGlyAsp 178

seq_name: p1r2:F86769

seq_documentation_block:
deoxyribose kinase [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 31-Mar-2001
C:Accession: F86769
R:Boilotin, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. in Press, 2001
A:Title: The complete genome sequence of the lactic acid bacterium.
A:Reference number: A86625
A:Accession: F86769
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-212 <STO>
A:Cross-references: GB:AF005176; NID:q12724122; PIDM:AAK05256.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: dukB
C:Superfamily: Lactobacillus acidophilus deoxyadenosine kinase

alignment_scores:
Quality: 141.00 Length: 181
Ratio: 1.382 Gaps: 8
Percent Similarity: 56.354 Percent Identity: 27.624

alignment_block:
US-09-416-579A-1 x F86769 ..
Align seq 1/1 to: F86769 from: 1 to: 212

```

```

67 GTCTCATCGAGGCAACATCGCAGCGGAGAGACCATGATTGACCA 116
   ::::: ||| ||| ||| ||| ||| |||
2 IleValLeuAlaGlyThrIleGlyAlaGlyLysSerSerLeuAlaLysAl 18
117 CTTC...GAGAAGTACAAAGACATTTGCTGCTGACCGACCCGCTGC 163
   ::| ||| ||| ||| ||| |||
18 aLeuGlyGlnHisLeuTyrThrAspAlaPhe..... 28
164 AGAAGTGGCGCAACGTCAACGGGTAAATCTGCTGAGAGCTGATGACAA 213
   ::::: ||| ||| ||| ||| ||| |||
29 ....TyrGlnAlaValAspAsnAsnProValLeuAspLeuTyrGln 43
214 GATCCCAAGAAAGTGGGCGCATGCCCTTTCAGAGTATGTACAGCTGACCAT 263
   ||| ||| ||| ||| ||| ||| |||
44 AspProGlnTyrTyrAlaPheLeuLeuGlnIleTyrPheLeuAsnLysAr 60
264 GCTGCAAGTGCACACCGCCCAACCAACAAGAGCTAAATTAARGGAGC 313
   ::| ||| ::| ||| ::| ||| ::| |||
60 gPheGlnSerIleLysMetAlaTyrArgGlnAspAsnValLeuAspA 77
314 GCTCATTTTAAAGCTGCTGCTATGCTTGTGGAGAACATGGCAGCAAC 363
   ||| ||| ||| ||| ||| ||| |||
77 rGSerIlePheGlnAspGlnLeuPheLeuThrLeuAsnTyr...LysAsn 92
364 GGCCTGCGTG.....GAGCAGCGCATGTAC.....AA 389
   ||| ::::: ||| ||| ||| ||| |||
93 GlnAsnValThrLysThrGlnLeuGlnIleTyrGlnAspLeuAlaAs 109
390 TACGCTGAGAGAGTGTACAAATTCATCGAAGAGCTCATTCACCTGCAGG 439
   ||| ||| ||| ||| ||| |||
109 MetLeuGlnGlu.....LeuAspGlyMetProLysLysArgp 122
440 CGGACCTCATCATATATCTGCGCACCTGCGGAGGTGGCGTACGAAAGC 489
   ||| ||| ||| ||| ||| ||| |||
122 roAspLeuLeuValTyrIleAspValSerPheGlnTyrMetLeuSerArg 138
490 ATCCGAGCGGCGCTCTTTCAGAGAGAGCTGC.....GTGCCGCT 530
   ||| ||| ||| ||| ||| ||| |||
139 IleAlaGlnArgGlyArgSerPheGlnGlnIleGlnAsnGlnGlnIle 155
531 TAAG...TACCTTCAGAGAGCTGCATGAGTTCACACGAGACTCG 570
   ||| ||| ||| ||| ||| ||| |||
155 uLysAspTyrTyrAlaGlnValHisAspGluTyrProAspTrp 169

seq_name: p1r1:K18BHS

seq_documentation_block:
thymidine kinase (EC 2.7.1.21) - salmireline herpesvirus 1 (strain 11(Onc))
C:Species: salmireline herpesvirus 1
A:Note: host Salmireline sciureus (common squirrel monkey)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 19-Jan-2001
C:Accession: A33374
R:Hones, R.W.; Craxton, M.A.; Williams, L.; Gompels, U.A.
J. Gen. Virol. 70, 3003-3013, 1989
A:Title: A comparative analysis of the sequence of the thymidine kinase gene of a gam
A:Reference number: A33374; M01D:90063548
A:Accession: A33374
A:Molecule type: DNA
A:Residues: 1-527 <HON>
A:Cross-references: EMBL:D00543; NID:g221853; PIDN:BA00432.1; PID:g221856
C:Superfamily: salmireline herpesvirus 1 thymidine kinase; herpesvirus thymidine kinas
C:Keywords: ATP; DNA biosynthesis; nucleotide binding; P-loop; phosphotransferase
F:209-492/Domain: herpesvirus thymidine kinase homology <Htk>
F:216-223/Region: nucleotide-binding motif A (P-loop)
F:311-315/Region: nucleotide-binding motif B
F:222/Binding site: ATP (Lys) #status predicted

alignment_scores:
Quality: 134.50 Length: 257
Ratio: 1.027 Gaps: 13
Percent Similarity: 50.973 Percent Identity: 25.292

```

alignment_block:

US-09-416-579A-1 x KIBHS

Align seg 1/1 to: KIBHS from: 1 to: 527

```

37 AACTAGCGGAGGCGACCCCTTACCGCTTCATGAGGAGCAACAT 86
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
202 LysHisGlnAspProValProAlaPheHisLeuGluGlySerI 218
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
87 CGGACCGGAGAGCAGCATTTG.....AACCTTCGAGAACT 127
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
218 egiValGlyLysThrThrLeuLeuSerMetAsnGlyLeuGly 235
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
128 ACAAGACGACATTTGCTGCTGAGCGAGCGCGAGAGAGGCGAAC 177
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
235 LysAsn...ValLeuAlaPheHisGluProIleAlaYrThrAsp 250
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
178 GTCAAGGCGTAAATCTGCTGAGCTGATGACAA..... 213
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
251 ValPheSer...AsnSerLeuGluValYrLysLeuThrLeuPro 266
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
214 .....GATCCCAAGA 223
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
266 alyValGlyAlaGlyThrSerAsnSerAlaLysLeuAlaCysGlu 283
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
224 AGTGGGCGCATGCCCTTTCAGACTTATGTCAGCGTCGACCATG 273
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
283 yspHeAlaSerProLeuLeuAlaLeuLysThrAlaThr..... 297
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
274 CACACCGCGCCCAACACAGACGTAATA.....ATAAR 308
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
298 LeuSerProLysAsnSerLeuLeuSerSerAspMetTrpValMet 314
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
309 GGAGGCTTCATTTTACCGCTGCTATGCTGCTGAGAGAACATG 358
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
314 eAspArgHisProLeuSerAlaThrValAlaPheProYrMetHis 331
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
359 GAACGCGCTGCTGAGAGCGGATGTACAAATACGCTGAGAGTGT 408
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
331 InAsnGlyPheLeu.....SerPheSerHisLeuIleGlnLeu 345
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
409 AAGTTCATCGAAGAGTCCATTCACCTCGAGCGGACCATCATAT 458
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
346 SerPheLysAlaSer.....ArgGlyAspAsnIleIleLeuLe 358
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
459 GCGACCTCGCGGAGGTGGCTGACGAACGATCCGCGAGCGGCTG 508
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
358 uAsnLeuAsnSerGlnGluAsnLeuLysArgValLysArgAsn 375
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
509 CTGAGGAGAGCTGCGCTGCGCTTAAGTACCTTCAGAGAGCTG 549
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
375 ysgGluLysSerValSerIleGlnHisIleArgLeuLeuAsn 391
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
550 ..CATGAGTTCACACGAGCTGTTGATGATCACACAGAG..CG 593
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
392 TyrHisAlaValYrCysAlaThrLeuLeuValGlnAsnPheThr 408
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
594 GTCTGCAAGGTCTAGTCTCGATGCC.....GATCTGAAC. 630
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
408 uGluIleValGluValCysPheAsnAlaLysHisIleThrAsp 425
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
631 .....CTGGAACAACTTGGACCGGAGTACCGCGC 660
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
425 eSerLysProSerPheLeuAlaLysHisValSerThrGlu... 440
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
661 TCGAGAGAGCAGCATATTTCAG 681
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
441 LeuLysSerSerIlePheAsn 447
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
seq_name: p1r2:B56168
seq_documentation_block:

```

deoxyguanosine kinase (EC 2.7.1.113) - Lactobacillus acidophilus
 N:Alternate names: dGuo kinase subunit
 C:Species: Lactobacillus acidophilus
 C:Date: 03-Oct-1995 #sequence_rev15ion 03-Oct-1995 #ext_change 18-Jun-1999
 C:Accession: B56168; D54257
 R:Ma, G.T.; Hong, Y.S.; Ives, D.H.
 J. Biol. Chem. 270, 6595-6601, 1995
 A:Title: Cloning and expression of the heterodimeric deoxyguanosine kinase/deoxyadeno
 A:Reference number: A56168; M01D:95204449
 A:Accession: B56168
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-224 <MA>
 A:Cross-references: GB:001881, NID:9403982; PIDN:AAB09751.1; PID:9403984
 A:Experimental source: strain R-26
 R:Ikeda, S.; Ma, G.T.; Ives, D.H.
 Biochemistry 33, 5328-5334, 1994
 A:Title: Heterodimeric deoxynucleoside kinases of Lactobacillus acidophilus R-26: fun
 A:Reference number: A54257; M01D:94227067
 A:Accession: D54257
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 2-18 <IK>
 A:Experimental source: R-26
 A:Note: sequence extracted from NCBI backbone (NCBIP:146751)
 C:Superfamily: Lactobacillus acidophilus deoxyadenosine kinase
 C:Keywords: phosphotransferase

alignment_scores:

Quality	Ratio	Length	Gaps
Percent Similarity: 60.849	Percent Identity: 22.642	212	8

alignment_block:

US-09-416-579A-1 x B56168

Align seg 1/1 to: B56168 from: 1 to: 224

```

67 GTCTCATTCGAGGCGAATCGCGACGGAGAGACCATATTGAAACA 116
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
4 ILeValLeuSerGlyProIleGlyAlaGlyLysSerSerLeuThrGly 20
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
117 CTTCGAGAACTACAGACGACATTTGCTGCTGAGACCGCGCTGACA 166
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
20 eLeuSerLysTrp.....LeuGlyThrAsnProPhe.... 30
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
167 AGTGGCGCAACGTCAACGGGGTAATCTGCTGAGAGCTGATGTAAGA 216
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
31 ..TyrGluSerValaAspAsnProValLeuProLeuPheYrGluAsn 46
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
217 CCAGAAAGTGGGCGCATGCCCTTTCAGAGTATGTCAGCGTACCAT 266
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
47 ProLysLysTrpAlaPheLeuLeuGlnValYrPheLeuAsnThrArg 63
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
267 GCAGTGCACACCGCCCAACCAAGAGCTAATAAATAAAGGAGCGCT 316
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
63 eArgSerIleLysSerAlaLeuThrAspAsnAsnValLeuAspArg 80
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
317 CCATTTTACGCGCTGCTATTCCTGCTGAGAGACATGCGAGAAAC 366
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
80 eTrpLysGluAspAlaLeuPhe.....GlnMetAsnAla 92
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
367 TCGCTGAGAGGCGCATGACAAATACGCTGAGAGAGTGTATGATTC 416
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
93 AspIleGlyAlaGlnAlaThrProGluGluValaAspThrTrpYrG 109
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
417 CGAAGAGTCCATT.....CACCTGCGAGCGGCGAC 445
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
109 uHisAsnMetMetSerGluLeuAspArgMetProLysLysAsnPro 126
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
446 TCATCATATATCGCGACACCTGCGGAGAGTGGCTGACGAACGATCG 495
   |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

```

126 euleuValHisIleaspValSerTyrAspThrMetLeuLysArgIleGln 142
496 CAGGGGGCTGCTGTGAGGAGAGTCGCGCTTACAGCTTCAGCA 545
143 LysArgIleGlnGlyMetTyrGln.....LeuSerTyrAspProth 156
546 GCTCATGAGTTCACGACGAGCTGTATACACAGACAGACCGCAGT 595
156 rleuGluAspTyrTyrLysArg...LeuLeuArgTyrTyrLysProThr 172
596 CGTCAGAGTC.....CTAGCTCGATGCCGATCTGAC 630
172 yralAlaLysTyrAspTyrSerProLysMetThrIleAspGlyAsp...Lys 187
631 CTGGAAAACATTTGGCACCAGTACACGCTCGAG 666
188 LeuAspMetAlaSerGluAspArgGln 199

```

seq_name: p1r1:G70129

seq_documentation_block:

deoxyguanosine/deoxyadenosine kinase I (EC 2.7.1.-) subunit 2 dck homolog - Lyme disease
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #extl_change 10-Sep-1999
C:Accession: G70129
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Keriavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
R:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MID:98065943
A:Accession: G70129
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
R:Molecule type: DNA
A:Residues: 1-205 <KLE>
A:Cross-references: GB:AE001134; GB:AE000783; NID:92688133; PIDN:AAC66630.1; PID:9268813
A:Experimental source: strain B31
C:Superfamily: Lactobacillus acidophilus deoxyadenosine kinase
C:Keywords: phosphotransferase

alignment_scores:
Quality: 132.50 Length: 196
Ratio: 1.380 Gaps: 9
Percent Similarity: 48.980 Percent Identity: 24.490

alignment_block:
US-09-416-579A-1 x G70129 ..

Align seg 1/1 to: G70129 from: 1 to: 205

```

67 GTCTTCATCGAGGCAACATCGCGGAGGGAAGACACGATTTGAC.. 114
.....|
7 rleuValIleGluGlyLeuIleGlyValGlyLysThrThrLeuGlyAsnI 23
115 .....CACTTCGAGAAGTACAGACGACA 139
23 eLeuSerLysGluLeuGluValProPheTyrSerGluLeuAsnAsp 40
140 TTTGCTGCTGACCGAGCCGTCGAGAAAGTGCGCAACGTCACGCGG 189
..|
40 hethrLeuAla..... 43
190 AATCTCTGAGAGTGATGTACAAAGATCCCAAGAGTGGCCATGCCCT 239
.....|
44 ...ValLeuAspLysPheTyrLysAspLysSerArgTyrPalaPhePro 59
240 TCAGAGTTATGTCACGCTGACCATGCTGCAGTGCACACCGCCCAACCA 289
|||
59 Igh.....IleAsnPheLeu.....A 65
290 ACAAGAACTAAATAATANAGAGCGCTCCATTTTAGCGCTGCTATTCG 339

```

```

|||.....|
65 snGluArgPheLysLeuIle...LysGlyValPheArgThr..... 77
340 TTCGTGAGAACATCCGACGAACCGCTGCTGAGCAGGCGCATGTAC.. 387
78 .....LysGlyLysIleLeuAspArgSerIleTyrG1 88
388 .....AATACGCTGAGAG 400
88 yAspCysValPheAlaSerLeuLeuAsnCysAspGlyHisIleSerAsp 105
401 AGTGTACAGTTCATCGAAGATCCATT.....CACCTG 435
105 LglutTyrLysIleTyrIleAspLeuLeuAspAsnMetLeuGluHisSer 121
436 CAG...GGGACCTCATCATATATCTGCGCACCTGCGCGAGGTGGGTA 482
122 GlnArgProSerLeuLeuValTyrLeuAspCysSerIleAspGluValG1 138
483 CGAAGCATCCGGCAGCGGGCTGTTCTGAGGAGAGCTGCGCCGCTTA 532
138 uArgArgIleLysAsnArgAsnArgSerPheGluMetAsnIleProArg 155
533 AATACCTCAGAGACTGCATGAGTTGCACCAAGACTGG 570
155 spTyrLeuGluGlyLeuAsnArgLysTyrLeuLysTrp 167

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 05:42:59 ; Search time 48.33 seconds

(Without alignments)
2949.540 Million cell updates/sec

Title: US-09-416-579a-1

Perfect score: 753

Sequence: 1 atgcgcgagcagcagcctctg.....gccagaggtccagataa 753

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 segs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCUTS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46.8	6.2	9551	1	US-08-056-200-93
2	46.8	6.2	9551	2	US-08-800-644-93
3	41.2	5.5	36519	3	US-08-923-137-2
4	39.2	5.2	467	2	US-08-476-176B-17
5	39.2	5.2	467	3	US-08-127-721A-17
6	39.2	5.2	467	3	US-08-485-246A-17
7	39.2	5.2	468	2	US-08-476-176B-13
8	39.2	5.2	468	3	US-08-127-721A-13
9	39.2	5.2	468	3	US-08-485-246A-13
10	38.2	5.1	467	2	US-08-476-176B-15
11	38.2	5.1	467	3	US-08-127-721A-15
12	38.2	5.1	467	3	US-08-485-246A-15
13	38.2	5.1	468	2	US-08-476-176B-11
14	38.2	5.1	468	3	US-08-127-721A-11
15	38.2	5.1	468	3	US-08-485-246A-11
16	38.2	5.1	7218	1	US-08-232-463-14
17	37	4.9	44377	2	US-08-804-227C-7
18	37	4.9	44377	2	US-08-804-198-1
19	36.4	4.8	831	1	US-07-882-329-1
20	36.4	4.8	831	1	US-08-183-213-1
21	36.4	4.8	1120	1	US-07-882-329-3
22	36.4	4.8	1120	1	US-08-183-213-3
23	36.4	4.8	1329	1	US-07-882-329-4
24	36.4	4.8	1329	1	US-08-183-213-4
25	36	4.8	477	4	US-08-998-416-132
26	36	4.8	2540	2	US-08-231-193A-15
27	36	4.8	2540	2	US-08-486-273A-15

28	36	4.8	2540	3	US-08-480-474-15	Sequence 15, Appl
29	36	4.8	2540	3	US-08-940-086A-15	Sequence 15, Appl
30	36	4.8	2621	3	US-08-436-332B-11	Sequence 1, Appl
31	36	4.8	2727	3	US-08-436-332B-13	Sequence 13, Appl
32	36	4.8	2835	3	US-08-436-332B-12	Sequence 12, Appl
33	36	4.8	2916	3	US-08-436-332B-11	Sequence 11, Appl
34	36	4.8	3007	2	US-08-231-193A-33	Sequence 33, Appl
35	36	4.8	3007	2	US-08-480-474-33	Sequence 33, Appl
36	36	4.8	3007	3	US-08-480-474-33	Sequence 33, Appl
37	36	4.8	3007	3	US-08-940-086A-33	Sequence 33, Appl
38	36	4.8	3070	2	US-08-231-193A-39	Sequence 39, Appl
39	36	4.8	3070	2	US-08-480-474-39	Sequence 39, Appl
40	36	4.8	3070	3	US-08-940-086A-39	Sequence 39, Appl
41	36	4.8	3070	3	US-08-480-474-39	Sequence 39, Appl
42	36	4.8	3155	2	US-08-231-193A-13	Sequence 13, Appl
43	36	4.8	3155	2	US-08-486-273A-13	Sequence 13, Appl
44	36	4.8	3155	3	US-08-480-474-13	Sequence 13, Appl
45	36	4.8	3155	3	US-08-940-086A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-08-056-200-93

Sequence 93, Application US/08056200

Patent No. 5616500

GENERAL INFORMATION:

APPLICANT: Steinert, Peter M.

APPLICANT: Lee, Seung-Chul

APPLICANT: Kim, In-Gyu

APPLICANT: Chung, Soo-Il

APPLICANT: Park, Sang-Chul

TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and

TITLE OF INVENTION: Methods of Using Same

NUMBER OF SEQUENCES: 117

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson & Bear

STREET: 620 Newport Center Drive, Sixteenth Floor

CITY: Newport Beach

STATE: CA

COUNTRY: U.S.A.

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/056,200

FILING DATE: 30-APR-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fredrick, Michael F.

REGISTRATION NUMBER: 36,799

REFERENCE/DOCKET NUMBER: NIH054,001A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (714) 760-0404

TELEFAX: (714) 760-9502

INFORMATION FOR SEQ ID NO: 93:

SEQUENCE CHARACTERISTICS:

LENGTH: 9551 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 1507..1644

FEATURE:

NAME/KEY: Intron

SEQUENCE CHARACTERISTICS:
LENGTH: 467 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 22..447
NAME/KEY: mat_peptide
LOCATION: 79..447
OTHER INFORMATION: /product="heavy chain variable
OTHER INFORMATION: region C21-Hay3"
US-08-127-721A-17

Query Match 5.2%; Score 39.2; DB 3; Length 467;
Best Local Similarity 52.4%; Pred. No. 0.016; Mismatches 78; Indels 0; Gaps 0;
Matches 86; Conservative 0;

QY 32 gggcccaagtagccggagggcaccgacctcaccgtccctcagagggcaacatcgga 91
DB 278 GGGCCACCTTCAACCCGACACCCAGCGCCAGCACCCTTACATGAGCTGAGCAGCTGA 337
QY 92 gcgaggagaccagctattgaaccacttcgagaagtaacaagaagacattgcctgtga 151
DB 338 GGAGCGAGGACACCGCGGTCTACTGCGCAGGTTCCAGCCCTTCAACCGCAGCAACT 397
QY 152 ccgagcccgctcgagaagtggcggaacgtcaacgggttaactcg 195
DB 398 ACGACTACTTCTGACTGAGGGCGGCGACGACCTGTGACCGTG 441

RESULT 6

US-08-485-246A-17
Sequence 17, Application US/08485246A
Patent No. 6072035
GENERAL INFORMATION:
APPLICANT: Hardman, No. 6072035man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: Immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6072035artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,246A
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6072035ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306

INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 22..447
NAME/KEY: mat_peptide
LOCATION: 79..447
OTHER INFORMATION: /product="heavy chain variable
OTHER INFORMATION: region C21-Hay3"
US-08-485-246A-17

Query Match 5.2%; Score 39.2; DB 3; Length 467;
Best Local Similarity 52.4%; Pred. No. 0.016; Mismatches 78; Indels 0; Gaps 0;
Matches 86; Conservative 0;

QY 32 ggaaccaagtagcgagggagccaccgacctcaccgtccctcagagggcaacatcgga 91
DB 278 GGGCCACCTTCAACCCGACACCCAGCGCCAGCACCCTTACATGAGCTGAGCAGCTGA 337
QY 92 gcgaggagaccagctattgaaccacttcgagaagtaacaagaagacattgcctgtga 151
DB 338 GGAGCGAGGACACCGCGGTCTACTGCGCAGGTTCCAGCCCTTCAACCGCAGCAACT 397
QY 152 ccgagcccgctcgagaagtggcggaacgtcaacgggttaactcg 195
DB 398 ACGACTACTTCTGACTGAGGGCGGCGACGACCTGTGACCGTG 441

RESULT 7

US-08-476-176B-13
Sequence 13, Application US/08476176B
Patent No. 5958708
GENERAL INFORMATION:
APPLICANT: Hardman, No. 5958708man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: Immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958708artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,176B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5958708ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110

FLYING DATE: 424
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127, 721
FLYING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952, 802
FLYING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6072035ak, Henry P.
REGISTRATION NUMBER: 33, 200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CI/P
TELECOMMUNICATION INFORMATION:
DATE:

Query Match	5.1%;	Score 38.2;	DB 2;	Length 467;
Best Local Similarity	52.1%;	Pred. No. 0.032;		
Matches 85; Conservative	0;	Mismatches 78;	Indels 0;	Gaps 0;

Qy	33	gaccaagatcgcgcgaaggagaccacccttaccgcgtctctatctgagaggaatctgagag	92
Db	279	GGCCACCTTACACCGCCGACGACGCGCCGACCCGCTTACATGACATGACAGCCTTAC	333
Qy	93	cggaagaccacglatltaaacacttcgagaagataaagaacatcttgcctctaac	155
Db	339	CAGGAGAGACACCGCGGTACTACTGCGCGCAGGTTACGCACTTACGGCGCAGCAACTA	396
Qy	153	cgagccgcctcgagaagctgcgaacgttaacaggggtataactctg	195
Db	389	CTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	441

Qy 153 cgagccgcctcgaagaatgycgcgaactcaacggygtaactctg 19
||| | |||| | ||| | ||| | ||| | ||| | ||| | |||
Db 399 CGACTACTTCGACTACTGGGGCCAGGGACACCTGTGACCGTC 44

RESULT 11
HS-08-127-

Sequence 15: Application US/08127721A
 Patent No. 6066718
 GENERAL INFORMATION:
 APPLICANT: Hardman, No. 6066718man
 APPLICANT: Kolbinger, Frank
 APPLICANT: Saldanha, Jose
 TITLE OF INVENTION: Reshaped monoclonal antibodies against an
 TITLE OF INVENTION: Immunglobulin isotype
 NUMBER OF SEQUENCES: 55
 CORRESPONDENCE ADDRESS:

CONTRACT NO. 6066718artis Patent and Trademark Department
ADDRESS: NO. 6066718 Route 10
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

CONDUCTED BY: 1
2
3 APPLICATION NUMBER: US/08/127,721A
4
5 FILING DATE: 27-SEPTEMBER-1993
6
7 CLASSIFICATION: 424
8
9 PRIOR APPLICATION DATA:
10
11 APPLICATION NUMBER: US 07/952,802
12
13 FILING DATE: 25-SEPTEMBER-1992
14
15 ATTORNEY/AGENT INFORMATION:
16
17 NAME: No. 6066718ak, Henry P.
18
19 REGISTRATION NUMBER: 33,200
20
21 REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIH
22
23 TELECOMMUNICATION INFORMATION:
24
25 TELEPHONE: (908) 277-5110
26
27

APPLICANT: Hardman, No. 6072035man
 APPLICANT: Koldinger, Frank
 APPLICANT: Saldana, Jose
 TITLE OF INVENTION: Reshaped monoclonal antibodies against an
 TITLE OF INVENTION: Immunoglobulin isotype
 NUMBER OF SEQUENCES: 55
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: No. 6072035artls Patent Department
 STREET: 59 Route 10
 City: East Hanover
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07936-1080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/485,246A
 FILING DATE:
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/127,721
 FILING DATE: 27-SEPTEMBER-1993
 APPLICATION NUMBER: US 07/952,802
 FILING DATE: 25-SEPTEMBER-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6072035ak, Henry P.
 REGISTRATION NUMBER: 33,200
 REFERENCE/DOCKET NUMBER: 4-19276/A/P2/C1P
 TELECOMMUNICATION INFORMATION:

Sequence 11, Application US/084/61/6B
Patent No. 5958708
GENERAL INFORMATION:
APPLICANT: Hardman, No. 5958708man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: Immunoglobulin Isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958708artle Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,176B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/955,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5958708ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP

TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 22..447
FEATURE:
NAME/KEY: mat-peptide
LOCATION: 79..447
OTHER INFORMATION: /product= "heavy chain variable
OTHER INFORMATION: region C21-H1"
US-08-476-176B-11

Query Match 5.1%; Score 38.2; DB 2; Length 468;
Best Local Similarity 52.1%; Pred. No. 0.032;
Matches 85; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Qy 33 gaccagaagtcgagagagcaccaccccttcacgcgtccatcgagggcaacatcgag 92
Db 279 GGCACACTTCCAGCCGACACCAACACCCGCTTACATGAGCTGAGCAGCTGAC 338
Qy 93 cgggaagaccagctatttgaaccacttcgagaagtagaagaagcagcatctgctgac 152
Db 339 CAGGAGGAGACACCGCGCTGACTACTGCGCCAGGTTACACCACTTCAGGGGAGCACTA 398
Qy 153 cgagcccgctcgagaagtgycgcaacgtlcaacggggttaactcg 195
Db 399 CGACTACTTGACTACTGCGGAGGACACCCCTGTGACCGTG 441

RESULT 14
US-08-127-721A-11
Sequence 11, Application US/08127721A
Patent No. 6066718
GENERAL INFORMATION:
APPLICANT: Hardman, No. 6066718man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6066718artis Patent and Trademark Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/127,721A
FILING DATE: 27-SEPTEMBER-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6066718ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 22..447
FEATURE:
NAME/KEY: mat-peptide
LOCATION: 79..447
OTHER INFORMATION: /product= "heavy chain variable
OTHER INFORMATION: region C21-H1"
US-08-127-721A-11

Query Match 5.1%; Score 38.2; DB 3; Length 468;
Best Local Similarity 52.1%; Pred. No. 0.032;
Matches 85; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Qy 33 gaccagaagtcgagagagcaccaccccttcacgcgtccatcgagggcaacatcgag 92
Db 279 GGCACACTTCCAGCCGACACCAACACCCGCTTACATGAGCTGAGCAGCTGAC 338
Qy 93 cgggaagaccagctatttgaaccacttcgagaagtagaagaagcagcatctgctgac 152
Db 339 CAGGAGGAGACACCGCGCTGACTACTGCGCCAGGTTACACCACTTCAGGGGAGCACTA 398
Qy 153 cgagcccgctcgagaagtgycgcaacgtlcaacggggttaactcg 195
Db 399 CGACTACTTGACTACTGCGGAGGACACCCCTGTGACCGTG 441

RESULT 15
US-08-485-246A-11
Sequence 11, Application US/08485246A
Patent No. 6072035
GENERAL INFORMATION:
APPLICANT: Hardman, No. 6072035man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6072035artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,246A
FILING DATE: 27-SEPTEMBER-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6072035ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP


```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908) 277-5110
: TELEFAX: (908) 277-4306
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 468 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 22..447
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 79..447
: OTHER INFORMATION: /product="heavy chain variable
: OTHER INFORMATION: region C21-H1"
US-08-485-246A-11

```

```

Query Match      5.1%; Score 38.2; DB 3; Length 468;
Best Local Similarity 52.1%; Pred. No. 0.032;
Matches 85; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

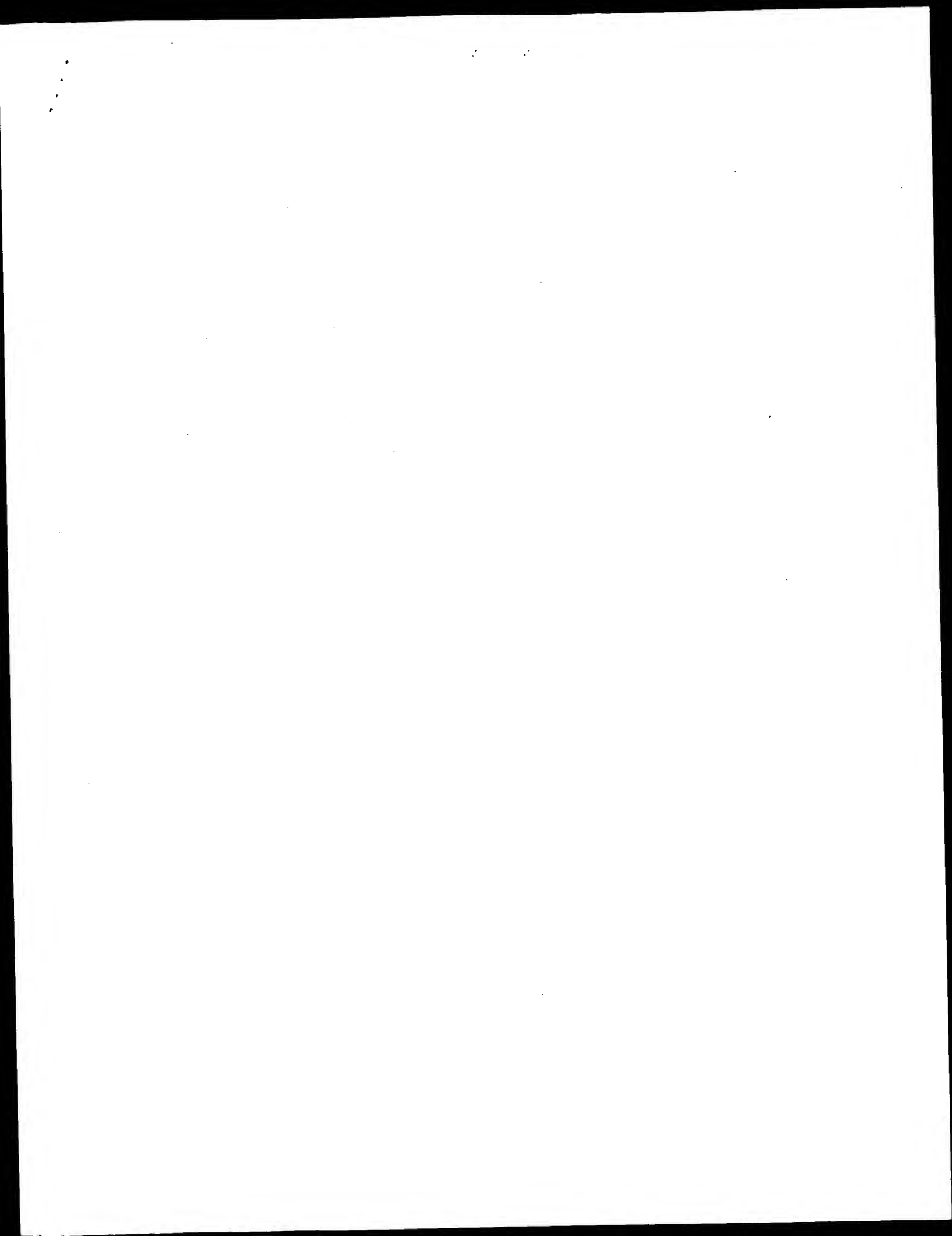
QY 33 gaccagaagcgcgagggcaccacccctcaccggtccatcgagggcaacatcgagcag 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 279 GCCACCTTACCGCCGACACCAACACCGCCCTACATGAGCTGAGCAGCCTGAC 338

QY 93 cgggaagaccacgatttgaccacatcgaggaagtacaagaagacatttgctctgac 152
   || || || || || || || || || || || || || || || || || || || ||
DB 339 CAGCGAGGACACCGCCGCTGTACTACTGCCGACAGTTCAGCCACTTCAAGCGGACCAACTA 398

QY 153 cgaagccgtcgagaagtgccgcaagtcacaagggtaaatctg 195
   || || || || || || || || || || || || || || || || || || || ||
DB 399 CGACTACTTGACTACTGGGGCCAGGGCACCCCTGTGTGACCGTG 441

```

Search completed: October 28, 2001, 06:24:06
 Job time: 2467 sec



[illegible]

XX The present sequence represents a Drosophila melanogaster
 CC deoxynucleoside kinase DNA sequence. The kinase remains stable
 CC during nucleoside monophosphate synthesis without the presence of
 CC stabilizing reagents. Deoxynucleoside kinases catalyze the
 CC phosphorylation of deoxynucleosides or deoxynucleosides. The enzyme
 CC therefore has an important role in the salvage pathway of nucleoside
 CC metabolism. The recombinant deoxynucleoside kinase is useful for
 CC the phosphorylation of a nucleoside to produce a nucleoside
 CC monophosphate.

XX Sequence 753 BP; 190 A; 215 C; 212 G; 135 T; 1 other;

Query Match 99.9%; Score 752.6; DB 21; Length 753;
 Best Local Similarity 100.0%; Pred. No. 4.4e-197;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggcgagagcagacatctctgcccgaaggggacacgaagcagcagccagcc 60
 DB 1 atggcgagagcagacatctctgcccgaaggggacacgaagcagcagccagcc 60
 QY ttccacctccatccaggggacacatcgccgagcggagagacacgaatttgaaacattc 120
 DB ttccacctccatccaggggacacatcgccgagcggagagacacgaatttgaaacattc 120
 QY 61 ttccacctccatccaggggacacatcgccgagcggagagacacgaatttgaaacattc 120
 DB 61 ttccacctccatccaggggacacatcgccgagcggagagacacgaatttgaaacattc 120
 QY 121 gagagatcaagaagacacatttgcctgctgacagcggccgctcgagaaagtggcgacgtc 180
 DB 121 gagagatcaagaagacacatttgcctgctgacagcggccgctcgagaaagtggcgacgtc 180
 QY 121 gagagatcaagaagacacatttgcctgctgacagcggccgctcgagaaagtggcgacgtc 180
 DB 121 gagagatcaagaagacacatttgcctgctgacagcggccgctcgagaaagtggcgacgtc 180
 QY 181 aacgggtaaatctgctggagctgattgataaagatcccaagaagtggccatccctt 240
 DB 181 aacgggtaaatctgctggagctgattgataaagatcccaagaagtggccatccctt 240
 QY 181 aacgggtaaatctgctggagctgattgataaagatcccaagaagtggccatccctt 240
 DB 181 aacgggtaaatctgctggagctgattgataaagatcccaagaagtggccatccctt 240
 QY 241 cagagttatgacagctgacacatctgcaagtcgacacacgcccacaaacaaagaagcta 300
 DB 241 cagagttatgacagctgacacatctgcaagtcgacacacgcccacaaacaaagaagcta 300
 QY 301 aaataaaggagcgtccatttttagcgctcgctattgcttgcgggagacatgagcaga 360
 DB 301 aaataaaggagcgtccatttttagcgctcgctattgcttgcgggagacatgagcaga 360
 QY 361 aacggctgcgtgagcagagcatgtacacatagctgagagagtggtacaaagtatcagaa 420
 DB 361 aacggctgcgtgagcagagcatgtacacatagctgagagagtggtacaaagtatcagaa 420
 QY 421 gagtcattcaactgacagcggaacctcaatatactgtgcacctgcgcgagagtggtcg 480
 DB 421 gagtcattcaactgacagcggaacctcaatatactgtgcacctgcgcgagagtggtcg 480
 QY 481 taggaacgcatccgagcagcggtcgttctgagagagacgtcgctgccccttaagtactt 540
 DB 481 taggaacgcatccgagcagcggtcgttctgagagagacgtcgctgccccttaagtactt 540
 QY 541 cagagagcagcatgatttcacacagagactggtgtatcacacagagagacagcagtcgtgc 600
 DB 541 cagagagcagcatgatttcacacagagactggtgtatcacacagagagacagcagtcgtgc 600
 QY 601 aaggtcctagtcctgcatgacgtatgaaacctggaacacatttgacagagttacacagcgc 660
 DB 601 aaggtcctagtcctgcatgacgtatgaaacctggaacacatttgacagagttacacagcgc 660
 QY 661 tcggagagcagcatatttgacagcgcattcgaagtaacacagcgccttcgcgttcgtgtg 720
 DB 661 tcggagagcagcatatttgacagcgcattcgaagtaacacagcgccttcgcgttcgtgtg 720
 QY 721 tcggagagcagcagcagagaggttcgcagataa 753
 DB 721 tcggagagcagcagcagagaggttcgcagataa 753

RESULT 2
 AAAL1980

ID AAA11980 standard; DNA; 792 BP.

XX AAA11980;

XX 07-AUG-2000 (first entry)

XX D. melanogaster deoxynucleoside kinase DM-dnk DNA.

XX Deoxynucleoside kinase; DM-dnk; fruit fly; nucleoside phosphorylation;
 XX nucleoside monophosphate synthesis; Insect; PCR; ds.

XX Drosophila melanogaster.

XX Location/Qualifiers

XX Key 22..774

XX CDS /*tag= a

XX /product= "Dm-dnk"

XX DE19846838-A1.

XX 13-APR-2000.

XX 12-OCT-1998; 98DE-1046838.

XX 12-OCT-1998; 98DE-1046838.

XX (HOFF) ROCHE DIAGNOSTICS GMBH.

XX Munch-Petersen B, Pliskur J, Sondergaard L, Ihlenfeldt H, Klein C;

XX Hagedorn U;

XX WPI; 2000-294113/26.

XX Recombinant Drosophila deoxynucleotide kinase useful for preparing

XX nucleoside monophosphates by phosphorylating nucleosides

XX Disclosure; Fig 5; 16pp; German.

XX This invention describes a novel recombinant kinase (I) obtainable from

XX insect cells, which is stable during nucleoside monophosphate synthesis

XX without the addition of stabilizing sulphydryl reagent or stabilizing

XX proteins and accepts all four natural deoxynucleotides. (I) is useful

XX for phosphorylating nucleosides to prepare nucleoside monophosphates.

XX Nucleoside monophosphates are useful as intermediates for nucleoside

XX triphosphates as polymerase chain reaction (PCR) reagents. This sequence

XX represents a DNA fragment which encodes the Drosophila melanogaster

XX deoxynucleoside kinase, designated DM-dnk, which is described in the

XX method of the invention.

XX Sequence 792 BP; 197 A; 224 C; 224 G; 146 T; 1 other;

XX Query Match 99.9%; Score 752.6; DB 21; Length 792;

XX Best Local Similarity 100.0%; Pred. No. 4.5e-197;

XX Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggcgagagcagacatctctgcccgaaggggacacgaagcagcagccagcc 60
 DB 1 atggcgagagcagacatctctgcccgaaggggacacgaagcagcagccagcc 60
 QY 61 ttccacctccatccaggggacacatcgccgagcggagagacacgaatttgaaacattc 120
 DB 61 ttccacctccatccaggggacacatcgccgagcggagagacacgaatttgaaacattc 120
 QY 121 gagagatcaagaagacacatttgcctgctgacagcggccgctcgagaaagtggcgacgtc 180
 DB 121 gagagatcaagaagacacatttgcctgctgacagcggccgctcgagaaagtggcgacgtc 180
 QY 142 gagagatcaagaagacacatttgcctgctgacagcggccgctcgagaaagtggcgacgtc 201
 DB 142 gagagatcaagaagacacatttgcctgctgacagcggccgctcgagaaagtggcgacgtc 201
 QY 181 aacgggtaaatctgctggagctgattgataaagatcccaagaagtggccatccctt 240
 DB 181 aacgggtaaatctgctggagctgattgataaagatcccaagaagtggccatccctt 240
 QY 202 aacgggtaaatctgctggagctgattgataaagatcccaagaagtggccatccctt 261
 DB 202 aacgggtaaatctgctggagctgattgataaagatcccaagaagtggccatccctt 261
 QY 241 cagagttatgacagctgacacatctgcaagtcgacacacgcccacaaacaaagaagcta 300

```

DB 262 cagagttatgtaacgctgacatgctgacgacacgcccacacacagaagcta 321
QY 301 aaaaataaragagcgtccatcttttagcgtcgctatctgctgagagaacatgcagca 360
DB 322 aaaaataaragagcgtccatcttttagcgtcgctatctgctgagagaacatgcagca 381
QY 361 aacggtcgtctgagacagagcgtacatacgtctgagaggtgtgtacaagtctgaa 420
DB 382 aacggtcgtctgagacagagcgtacatacgtctgagaggtgtgtacaagtctgaa 441
QY 421 ggttcattacactgcagagcggaacctcatatattctgcgaacctgcgcgaggtgag 480
DB 442 ggttcattacactgcagagcggaacctcatatattctgcgaacctgcgcgaggtgag 501
QY 481 tacgaacgcatccgagcagcggtctgtctgagagagctgcgtgcgcttaagtaacct 540
DB 502 tacgaacgcatccgagcagcggtctgtctgagagagctgcgtgcgcttaagtaacct 561
QY 541 caggaagctgcatgagttgacacaggaactggtgtatatacagcagagcagcagtcgtgc 600
DB 562 caggaagctgcatgagttgacacaggaactggtgtatatacagcagagcagcagtcgtgc 621
QY 601 aaggtccctagctctgcatgagcgtatctgaacctggaacattggcaaccgagttacagcgc 660
DB 622 aaggtccctagctctgcatgagcgtatctgaacctggaacattggcaaccgagttacagcgc 681
QY 661 tcgagagcagacataattcgacacatctcctcaagtaacacagcgcctgcgcgttcgtgtg 720
DB 682 tcgagagcagacataattcgacacatctcctcaagtaacacagcgcctgcgcgttcgtgtg 741
QY 721 tcgcccagacagcgccagagaggtgcgcagataa 753
DB 742 tcgcccagacagcgccagagaggtgcgcagataa 774

RESULT 3
AAAS1375
ID AAAS1375 standard; cDNA; 1113 BP.
AC AAAS1375;
AT 26-SEP-2000 (first entry)
DE D. melanogaster dNK, multisubstrate deoxyribonucleosidase cDNA.
KW dNK; multisubstrate deoxyribonucleosidase; kinase; cytosolic;
OS Drosophila melanogaster.
FH Key Location/Qualifiers
FT CDS 113..865
FT /tag- a
FT /product- Multisubstrate_deoxyribonucleosidase
PN NM0200036099-A1.
PD 22-JUN-2000.
PE 10-DEC-1999; 99MO-SE02314.
PR 11-DEC-1998; 98SE-0004298.
PA (KARL/) KARLSSON A.
PI Karlsson A, Johansson M;
PI
DR MPI: 2000-431584/37.
DR P-PSDB; AAY96812.
XX
XX New nucleic acid encoding a multisubstrate deoxyribonucleosidase with
PT homology to a Drosophila melanogaster sequence for activating

```

```

PT anti-viral and anti-cancer drugs to treat viral infections and cancer
XX
XX Claim 1: Page 33; 38pp; English.
PS
PS This cDNA encodes Drosophila melanogaster multisubstrate
XX deoxyribonucleosidase (dNK). The multisubstrate deoxyribonucleoside
CC changes nucleoside analogue products into active anti-cancer and
CC anti-viral drugs by phosphorylation, allowing cancer and viral
CC infections to be treated. Drosophila melanogaster deoxyribonucleoside
CC kinase is (unlike human deoxyribonucleoside kinase) a multisubstrate
CC enzyme and it catalyzes phosphorylation of pyrimidine and purine
CC deoxyribonucleosides. The catalytic rates of deoxyribonucleoside and
CC nucleoside analogue phosphorylation are 10 to 100 fold higher than the
CC maximal catalytic rates reported for mammalian enzymes, making it more
XX useful for activating anti-viral and anti-cancer drugs.
SQ
SQ Sequence 1113 BP; 324 A; 287 C; 272 G; 230 T; 0 other;

Query Match 98.3%; Score 740.2; DB 21; Length 1113;
Best Local Similarity 98.9%; Pred. No. 1,3e-193;
Matches 745; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 atgacgagagcagcatctctgcccgaagaaggagacgaatgacgacgagcaccgcc 60
DB 113 atgacgagagcagcatctctgcccgaagaaggagacgaatgacgacgagcaccgcc 172
QY 61 ttaccgtctcattcagagagcgaacatcgacgagcgggaagacacgatttgaccacctc 120
DB 173 ttaccgtctcattcagagagcgaacatcgacgagcgggaagacacgatttgaccacctc 232
QY 121 gaaagttacagaagaacacatttgcctgctgacccagacccgtcgagaagtggcgaagtc 180
DB 233 gaaagttacagaagaacacatttgcctgctgacccagacccgtcgagaagtggcgaagtc 292
QY 181 aacggggtaaatctgcttgagctgagctgatacaagaagatcccaagaatgggacatgcctt 240
DB 293 aacggggtaaatctgcttgagctgagctgatacaagaagatcccaagaatgggacatgcctt 352
QY 241 cagagttatgtaacgctgacacatgctgacgctgcacacccgcccaacaaagaagcta 300
DB 353 cagagttatgtaacgctgacacatgctgacgctgcacacccgcccaacaaagaagcta 412
QY 301 aaaaataaragagcgtccatcttttagcgtcgctatctgctgagagaactgcagca 360
DB 413 aaaaataaragagcgtccatcttttagcgtcgctatctgctgagagaactgcagca 472
QY 361 aacggtcgtctgagacagagcgtacatacgtctgagaggtgtgtacaagtctgaa 420
DB 473 aacggtcgtctgagacagagcgtacatacgtctgagaggtgtgtacaagtctgaa 532
QY 421 ggttcattacactgcagagcggaacctcatatattctgcgaacctgcgcgaggtgag 480
DB 533 ggttcattacactgcagagcggaacctcatatattctgcgaacctgcgcgaggtgag 592
QY 481 tacgaacgcatccgagcagcggtctgtctgagagagctgtgtgcgcttaagtaacct 540
DB 593 tacgaacgcatccgagcagcggtctgtctgagagagctgtgtgcgcttaagtaacct 652
QY 541 caggaagctgcatgagttgacacaggaactggtgtatatacagcagagcagcagtcgtgc 600
DB 653 caggaagctgcatgagttgacacaggaactggtgtatatacagcagagcagcagtcgtgc 712
QY 601 aaggtccctagctctgcatgagcgtatctgaaccttggaacattggcaaccgagttacagcgc 660
DB 713 aaggtccctagctctgcatgagcgtatctgaaccttggaacattggcaaccgagttacagcgc 772
QY 661 tcgagagcagacataattcgacgacatctcctcaagtaacacagcgcctgcgcgttcgtgtg 720
DB 773 tcgagagcagacataattcgacgacatctcctcaagtaacacagcgcctgcgcgttcgtgtg 832
QY 721 tcgcccagacagcgccagagaggtgcgcagataa 753

```


PD	09-OCT-1997.
PP	
PF	27-MAR-1997; 97MO-N200040.
PG	
PH	29-MAR-1996; 96NZ-0286284.
PI	(UYOT-) UNIV OTAGO.
PJ	
PK	Lyttle DJ, Robinson AJ;
PL	
PM	WPI: 1997-503116/46.
PN	
PO	Parapoxvirus vectors containing exogenous DNA - comprise parapox
PP	virus, especially orf virus, and exogenous DNA encoding a
PQ	polypeptide, e.g. an antigen useful in vaccines
PR	
PS	Claim 12: Fig 7; 73pp; English.
PT	
PV	
PW	
PX	
PY	
PZ	
QA	
QB	
QC	
QD	
QE	
QF	
QG	
QH	
QI	
QJ	
QK	
QL	
QM	
QN	
QO	
QP	
QQ	
QR	
QS	
QT	
QU	
QV	
QW	
QX	
QY	
QZ	
RA	
RB	
RC	
RD	
RE	
RF	
RG	
RH	
RI	
RJ	
RK	
RL	
RM	
RN	
RO	
RP	
RQ	
RR	
RS	
RT	
RU	
RV	
RW	
RX	
RY	
RZ	
SA	
SB	
SC	
SD	
SE	
SF	
SG	
SH	
SI	
SJ	
SK	
SL	
SM	
SN	
SO	
SP	
SQ	
SR	
SS	
ST	
SV	
SW	
SX	
SY	
SZ	
TA	
TB	
TC	
TD	
TE	
TF	
TF	
TH	
TI	
TJ	
TK	
TL	
TM	
TN	
TO	
TP	
TP	
TQ	
TR	
TS	
TT	
TU	
TU	
TV	
TV	
TW	
TX	
TX	
TY	
TY	
TZ	
TA	
TB	
TC	
TD	
TE	
TE	
TF	
TF	
TH	
TI	
TJ	
TK	
TL	
TM	
TN	
TO	
TP	
TP	
TQ	
TR	
TS	
TT	
TU	
TU	
TV	
TV	
TW	
TX	
TX	
TY	
TY	
TZ	
TA	
TB	
TC	
TD	
TE	
TE	
TF	
TF	
TH	
TI	
TJ	
TK	
TL	
TM	
TN	
TO	
TP	
TP	
TQ	
TR	
TS	
TT	
TU	
TU	
TV	
TV	
TW	
TX	
TX	
TY	
TY	
TZ	
TA	
TB	
TC	
TD	
TE	
TE	
TF	
TF	
TH	
TI	
TJ	
TK	
TL	
TM	
TN	
TO	
TP	
TP	
TQ	
TR	
TS	
TT	
TU	
TU	
TV	
TV	
TW	
TX	
TX	
TY	
TY	
TZ	
TA	
TB	
TC	
TD	
TE	
TE	
TF	
TF	
TH	
TI	
TJ	
TK	
TL	
TM	
TN	
TO	
TP	
TP	
TQ	
TR	
TS	
TT	
TU	
TU	
TV	
TV	
TW	
TX	
TX	
TY	
TY	
TZ	
TA	
TB	
TC	
TD	
TE	
TE	
TF	
TF	
TH	
TI	
TJ	
TK	
TL	
TM	
TN	
TO	
TP	
TP	
TQ	
TR	
TS	
TT	
TU	


```

FT      CDS /product= "Codon optimised gag"
FT      2613..5624
FT      /tag= 9
FT      /product= "Codon optimised pol"
FT      5710..6469
FT      /tag= h
FT      /label= Human_beta globin_sequence
FT      5897..5921
FT      polyA_site
FT      CDS /tag= 1
FT      6709..7569
FT      /tag= j
FT      /product= "beta-lactamase (bla)"
FT      8796..8908
FT      rep_origin
FT      /tag= k
FT      /label= SV40_origin_of_replication
FT      2583..2819
FT      /tag= 1
FT      /note= "This region consists of the gagpol overlap
FT      and cis-acting signals necessary for translation of
FT      pol which are not optimised"

```

```

PN      WO200015819-A1.
PD      23-MAR-2000.
PE      10-SEP-1999; 99WO-US20675.
PR      11-SEP-1998; 98US-0100022.
PR      12-SEP-1998; 98US-0100063.
PA      (CHIL-) CHILDRENS MEDICAL CENT.
PI      Gray JT, Mulligan RC;
DR      WPI: 2000-271455/23.
DR      P-PSDB; AAY70599, AAY70602.
XX      New packaging cell line for producing a viral accessory protein
XX      independent HIV derived retroviral vector particles, useful in gene
XX      therapy or gene replacement -
XX      Disclosure; Fig 10; 62pp; English.
XX
XX      The patent discloses new packaging cell line for producing a viral
XX      accessory protein independent lentivirus, preferably
XX      human immunodeficiency virus (HIV), derived retroviral vector particles.
XX      The packaging cell line comprises a mammalian cell, a retroviral DNA
XX      comprising a coding sequence for a lentivirus, preferably HIV, gagpol,
XX      where the coding sequence has been mutagenised to improve expression of
XX      the viral gagpol proteins, a second retroviral nucleotide sequence
XX      comprising the coding sequence for a heterologous envelope protein and a
XX      third retroviral nucleotide sequence comprising a DNA sequence of
XX      interest and lentivirus, preferably HIV, cis-acting sequences required
XX      for packaging, reverse transcription and integration.
XX      The packaging cell lines and viral particles can be used for gene
XX      therapy or gene replacement with improved safety. They can also be used
XX      in the development and production of vaccines and biochemical reagents.
XX      The present DNA is that of a packaging construct pMDMgpm2 which
XX      comprises a codon optimised HIV gagpol. The plasmid is defective for the
XX      production of the viral envelope and accessory proteins e.g. tat, vif,
XX      vpr, rev and Rev response element (RRE). It also lacks the viral
XX      sequences which are transcribed into mRNA, e.g. constitutive transport
XX      cell line.
XX
XX      Sequence 8908 BP; 2134 A; 2579 C; 2355 G; 1840 T; 0 other;

```

```

Query Match      5.4%; Score 41; DB 21; Length 8908;
Best Local Similarity 49.8%; Pred. No. 0.18;
Matches 104; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY      2 tggcgagagcagcatctgtgccgaagaaggaggaacgaatgacgagggaccacgcacct 61

```

```

DB      3614 tggcgagagagcactcctgagccgtgacgacatccagaagctgtgtggcaagctgaact 3873
QY      62 tcaccgctccatcgaagggcaacatcgcgagcggaagaccacgatttgaaccacttg 121
DB      3874 gggctcccgatcattacgacgacatcaaatcgccagctgtcaagctgtgtcgcgaca 3933
QY      122 agaagttacaagaacacatttgcctgtacccgagcccgctcgaagaattgctgcgaacgtca 181
DB      3934 ccaagccctgaccgaggtgtgtgcccctgacgagagagcgccgagctgtgcccgcaga 3993
QY      182 acggggtaaatctgtcgtgagctgtatgtac 210
DB      3994 accgagatcctcgaagagcccgctgcac 4022

RESULT 9
AAD02031
ID      AAD02031 standard; cDNA; 1180 BP.
XX
AC      AAD02031;
XX
DT      26-MAR-2001 (first entry)
XX
DE      Producing activating enzyme Carboxypeptidase G2 (CPG2) variant cDNA.
XX
KW      Carboxypeptidase G2; CPG2; gene directed enzyme producing therapy; GDEPT;
KW      glycosylphosphatidylinositol; GPI; cancer; therapy; mutant; mutein;
KW      CPG2(Q3) variant; ss.
XX
OS      Bacteria.
XX
XX      Synthetic.
XX
FH      Key Location/Qualifiers
FT      CDS 1..1176
FT      /tag= a
FT      /product= "CPG2(Q3) variant protein"
FT      /note= "CDS does not include start codon and has an
FT      additional stop codon at positions 1177-1179"
FT      /partial
XX
PN      WO200066752-A2.
XX
PD      09-NOV-2000.
XX
PF      28-APR-2000; 2000WO-GB01640.
XX
PR      01-MAY-1999; 99GB-0010077.
XX
PA      (ASTR ) ASTRAZENECA AB.
PA      (UYMA-) UNIV VICTORIA MANCHESTER.
PI      Castro MG, Emery SC, Lowenstein PR;
XX
XX      WPI: 2001-015983/02.
XX      P-PSDB; AAY71932.
XX
XX      Gene directed enzyme producing therapy using post translational
XX      glycosylphosphatidylinositol addition to producing activating enzyme to
XX      enable anchorage of enzyme at cell surface for cancer therapy -
XX
XX      Example 1d; Page 56; 60pp; English.
XX
XX      The present invention relates to a gene directed enzyme producing
XX      therapy (GDEPT) using post translational glycosylphosphatidylinositol
XX      (GPI) addition to a producing activating enzyme which enables anchorage
XX      of the enzyme at the cell surface. Carboxypeptidase G2 (CPG2) is a
XX      preferred producing activating enzyme. The invention also relates to an
XX      expression vector for expression of a GPI enzyme hybrid capable of
XX      anchorage to the surface of a mammalian cell. The expression vector
XX      comprises polynucleotide sequences encoding a signal peptide, an enzyme
XX      capable of activating a producing, and a post-translational GPI addition
XX      motif. The expression vector is useful in the manufacture of a medicament

```


AAV72046
ID AAV72046 standard; DNA; 1929 BP.
AC AAV72046;
XX
XX
DT 10-MAY-1999 (first entry)
XX
XX
DE Clone pNG4/ASB7VH-IGG2CH1/CPG2 R6 DNA.
XX
XX
XX Conjugate; cell targeting; cytotoxic drug; produg therapy system;
XX produg-converting enzyme; cell surface antigen; treatment; cancer;
XX inflammation; rheumatoid arthritis; antibody; ss.
XX
XX Synthetic.
XX
XX MO9851787-A2.
XX
XX 19-NOV-1998.
XX
XX
XX 05-MAY-1998; 98MO-GB01294.
XX
XX 10-MAY-1997; 97GB-0009421.
XX
XX (ZENE) ZENECA LTD.
XX
XX Blakey DC, Emery SC;
XX
XX WPI; 1999-059700/05.
XX
XX P-PSDB; AAM82739.
XX
XX
XX New gene construct expressing conjugate of targeting agent and
XX produg-converting enzyme - useful for, e.g. targeted production of
XX cytotoxic drug in vivo, especially for treatment of cancer
XX
XX Example 1e; Page 67-68; 100pp; English.
XX
XX
XX This sequence is used in a method for obtaining a novel gene construct
XX (A) which expresses, in cells of a mammal, a conjugate (B) of a
XX cell-targeting group (I) and a heterologous produg-converting enzyme
XX (II), and (B) is directed to leave the cell for selective localisation at
XX a cell surface antigen (Ag) recognised by (I). Delivery of (A) to a
XX target site, then administration of (III) is used for targeted release of
XX cytotoxic drug, specifically for treating cancer but also inflammation
XX such as rheumatoid arthritis. In situ generation of the targeting
XX antibody increases selectivity, reducing side effects at normal tissue.
XX The method is applicable to any antibody-directed enzyme produg therapy
XX system.
XX
XX Sequence 1929 BP; 409 A; 603 C; 597 G; 320 T; 0 other;
XX
XX

Query Match 5.3%; Score 39.8; DB 20; Length 1929;
Best Local Similarity 47.0%; Pred. No. 0.22;
Matches 119; Conservative 1; Mismatches 133; Indels 0; Gaps 0;
OY 150 gaccgaagccgcgcgaagtgcgcaacgcaagggtaaatctgctgagctgata 209
DB 1383 gcccgagcggcgcggaacgcgtgctgagagcttcgcgaactcgtgctgcgaacgatgaa 1442
OY 210 caaagatcccaagaagtggccatgccccttcagaagtatttcacgctgacacatgctga 269
DB 1443 catcgacgaagaagcaagcaactgcttcagtgccatgcgcaagggcgccaggt 1502
OY 270 gtcgcacacggcccaacaaagaagtaaaataargagacgctccattttaagc 329
DB 1503 ctggaacatcatcccgccagcgccacgctgacgacgcaagtgctacgagcgcaagga 1562
OY 330 tcgctattctcgttgagaaacatgcaagcaacgctgcgtgagcaagcgatgtacaa 389
DB 1563 ggaacttcgacgacgacatgaagacgtggaagaagcgcgacgacagaagaagctgcccga 1622
OY 390 tacgctgagagag 402

DB 1623 ggcgcagctgaag 1635
RESULT 14
ID AAV72075
ID AAV72075 standard; DNA; 1998 BP.
AC AAV72075;
XX
XX
XX 10-MAY-1999 (first entry)
XX
XX
XX Fusion protein (806.077 scFv-CPG2)2 DNA.
XX
XX
XX Conjugate; cell targeting; cytotoxic drug; plasmid; fusion protein;
XX produg-converting enzyme; cell surface antigen; treatment; cancer;
XX inflammation; rheumatoid arthritis; antibody; produg therapy system; ss.
XX
XX Synthetic.
XX
XX MO9851787-A2.
XX
XX 19-NOV-1998.
XX
XX 05-MAY-1998; 98MO-GB01294.
XX
XX 10-MAY-1997; 97GB-0009421.
XX
XX (ZENE) ZENECA LTD.
XX
XX Blakey DC, Emery SC;
XX
XX WPI; 1999-059700/05.
XX
XX P-PSDB; AAM82745.
XX
XX
XX New gene construct expressing conjugate of targeting agent and
XX produg-converting enzyme - useful for, e.g. targeted production of
XX cytotoxic drug in vivo, especially for treatment of cancer
XX
XX Example 17; Page 87; 100pp; English.
XX
XX
XX This sequence is used in a method for obtaining a novel gene construct
XX (A) which expresses, in cells of a mammal, a conjugate (B) of a
XX cell-targeting group (I) and a heterologous produg-converting enzyme
XX (II), and (B) is directed to leave the cell for selective localisation
XX at a cell surface antigen (Ag) recognised by (I). Delivery of (A) to a
XX target site, then administration of (III) is used for targeted release of
XX cytotoxic drug, specifically for treating cancer but also inflammation
XX such as rheumatoid arthritis. In situ generation of the targeting
XX antibody increases selectivity, reducing side effects at normal tissue.
XX The method is applicable to any antibody-directed enzyme produg therapy
XX system.
XX
XX Sequence 1998 BP; 420 A; 601 C; 628 G; 349 T; 0 other;
XX
XX

Query Match 5.3%; Score 39.8; DB 20; Length 1998;
Best Local Similarity 47.0%; Pred. No. 0.23;
Matches 119; Conservative 1; Mismatches 133; Indels 0; Gaps 0;
OY 150 gaccgaagccgcgcgaagtgcgcaacgcaagggtaaatctgctgagctgata 209
DB 1452 gcccgagcggcgcggaacgcgtgctgagagcttcgcgaactcgtgctgcgaacgatgaa 1511
OY 210 caaagatcccaagaagtggccatgccccttcagaagtatttcacgctgacacatgctga 269
DB 1512 catcgacgaagaagcaagcaactgcttcagatgtaacatgcgcaagcgccaggt 1571
OY 270 gtcgcacacggcccaacaaagaagtaaaataargagacgctccattttaagc 329
DB 1572 ctggaacatcatcccgccagcgccacgctgacgacgcaagtgctacgagcgcaagga 1631
OY 330 tcgctattctcgttgagaaacatgcaagcaacgctgcgtgagcaagcgatgtacaa 389

Db 1632 ggaactcgaagccgcacatgaagacgtggaagagcgcgacgacgaagaagctgcccga 1691
 QY 390 tacgctggaagag 402
 Db 1692 ggcgcgacgtgaag 1704

RESULT 15

AAV72059
 ID AAV72059 standard; DNA: 2019 BP.

AC AAV72059;

DT 10-MAY-1999 (first entry)

DE Plasmid pNG4/55.1scfv/CPG2 R6 DNA.

KW Conjugate; cell targeting; cytotoxic drug; prodrug-converting enzyme;
 cell surface antigen; treatment; cancer; inflammation; antibody;
 rheumatoid arthritis; prodrug therapy system; ss.

OS Synthetic.

PN WO9851787-A2.

PD 19-NOV-1998.

PE 05-MAY-1998; 98WO-GB01294.

PR 10-MAY-1997; 97GB-0009421.

PA (ZENE) ZENECA LTD.

PI Blakey DC, Emery SC;

DR WPI; 1999-059700/05.

DR P-PSDB; AAW82742.

PT New gene construct expressing conjugate of targeting agent and
 prodrug-converting enzyme - useful for, e.g. targeted production of
 cytotoxic drug in vivo, especially for treatment of cancer

PS Example 14; Page 77; 100pp; English.

CC This sequence is used in a method for obtaining a novel gene construct
 (A) which expresses, in cells of a mammal, a conjugate (B) of a
 cell-targeting group (I) and a heterologous prodrug-converting enzyme
 (II), and (B) is directed to leave the cell for selective localisation at
 a cell surface antigen (Ag) recognised by (I). Delivery of (A) to a
 target site, then administration of (II) is used for targeted release of
 cytotoxic drug, specifically for treating cancer but also inflammation
 such as rheumatoid arthritis. In situ generation of the targeting
 CC antibody increases selectivity, reducing side effects at normal tissue.
 CC The method is applicable to any antibody-directed enzyme prodrug therapy
 CC system.

SQ Sequence 2019 BP; 437 A; 604 C; 637 G; 341 T; 0 other;

Query Match 5.3%; Score 39.8; DB 20; Length 2019;
 Best local similarity 47.0%; Pred. No. 0.23;

Matches 119; Conservative 1; Mismatches 133; Indels 0; Gaps 0;

QY 150 gaccgagccgcgtcgagaagtgcgcaacgtcaacgggttaattgttgagctgatgta 209

Db 1473 gcccgagctggcgctggaacgctgtgtagagcttcgacctgctgcgcacagatgaa 1532

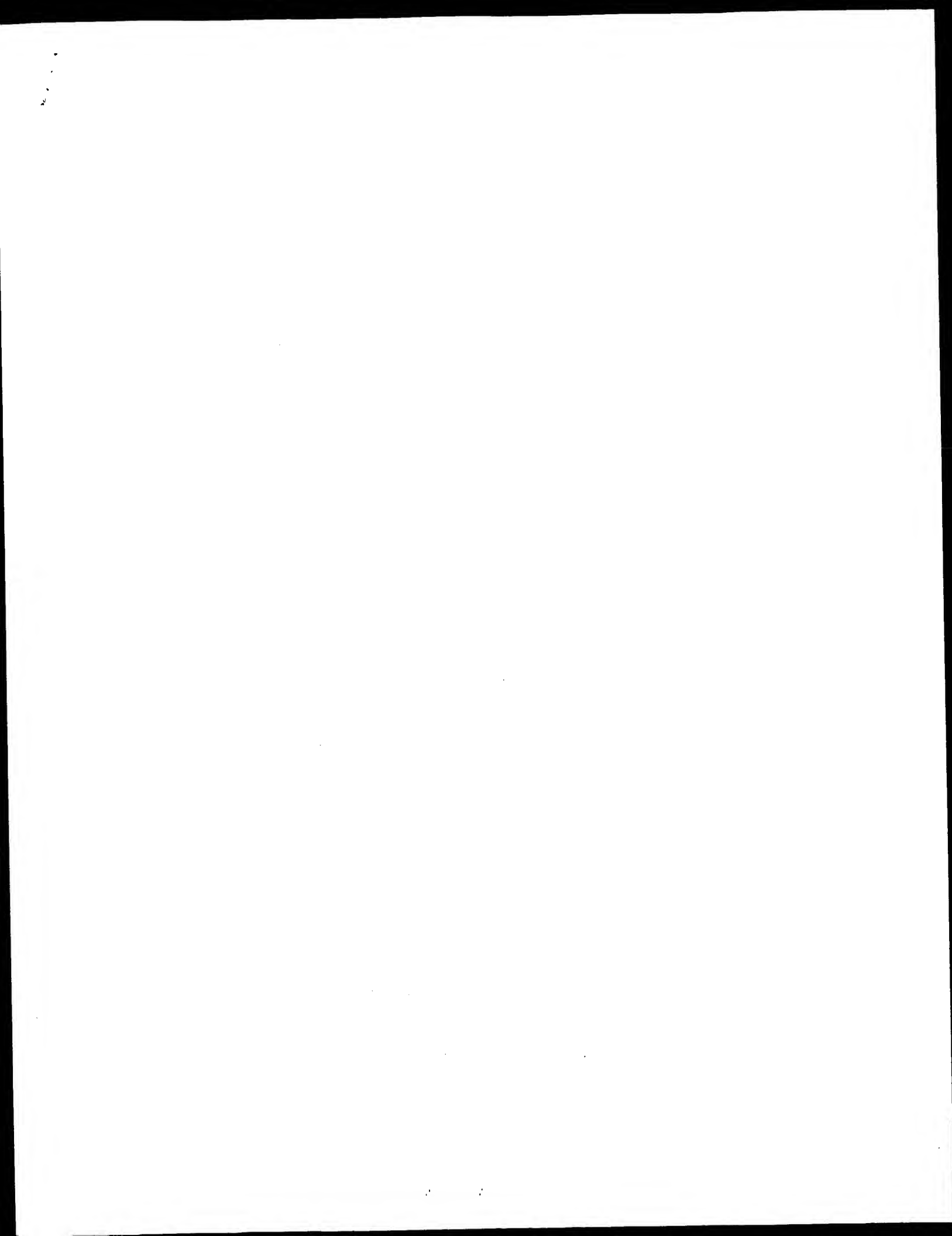
QY 210 caaagatcccaagaagtggccatgaccttcagagtatgtcacgctgacatgctgca 269

Db 1533 catcgacgacgaagcgcaagaccttccagcttgacatgcgcaagcgccgacaggt 1592

QY 270 gtgcgacacgcgcccaacacgaagctaaataatgagagctccatttttagcgc 329

Db 1593 ctggaacatcatcccccagccagcgccacgtgaaagcgagctgctacgcgcgcaacga 1652
 QY 330 tcgctattcttcgtggggaacatgcaacgagctcgtgagcagggcatgtacaa 389
 Db 1653 ggaactcgacgcccacatgaaagacgttgaaagagcgcgcgacgaagaagctgcccga 1712
 QY 390 tacgctggaagag 402
 Db 1713 ggcgcgacgtgaag 1725

Search completed: October 28, 2001, 06:26:04
 Job time: 2565 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 04:34:04 ; Search time 1077.11 Seconds
(without alignments)
10813.394 Million cell updates/sec

Title: US-09-416-579a-1

Perfect score: 753
Sequence: 1 atggcggaggcagcaccctg.....gccagagggtccgacagataa 753

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 segs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_ba3:*
- 4: gb_in1:*
- 5: gb_in2:*
- 6: gb_in3:*
- 7: gb_om:*
- 8: gb_om:*
- 9: gb_pat1:*
- 10: gb_pat2:*
- 11: gb_ph:*
- 12: gb_p11:*
- 13: gb_p12:*
- 14: gb_p13:*
- 15: gb_p14:*
- 16: em_ba1:*
- 17: em_ba2:*
- 18: em_fun:*
- 19: em_higo_hum:*
- 20: em_higo_inv:*
- 21: em_higo_rod:*
- 22: em_htg_hum1:*
- 23: em_htg_hum2:*
- 24: em_htg_hum3:*
- 25: em_htg_hum4:*
- 26: em_htg_hum5:*
- 27: em_htg_hum6:*
- 28: em_htg_hum7:*
- 29: em_htg_hum8:*
- 30: em_htg_inv1:*
- 31: em_htg_inv2:*
- 32: em_htg_other:*
- 33: em_htg_rod:*
- 34: em_hum1:*
- 35: em_hum2:*
- 36: em_hum3:*
- 37: em_hum4:*
- 38: em_hum5:*
- 39: em_hum6:*
- 40: em_hum7:*
- 41: em_in:*
- 42: em_om:*
- 43: em_or:*

44: em_ov:*

45: em_pat:*

46: em_ph:*

47: em_pl:*

48: em_to:*

49: em_sts:*

50: em_sy:*

51: em_un:*

52: em_v1:*

53: gb_sts1:*

54: gb_sts2:*

55: gb_sts3:*

56: gb_sy:*

57: gb_un:*

58: gb_v11:*

59: gb_v12:*

60: gb_htg1:*

61: gb_htg2:*

62: gb_htg3:*

63: gb_htg4:*

64: gb_htg5:*

65: gb_htg6:*

66: gb_htg7:*

67: gb_htg8:*

68: gb_htg9:*

69: gb_htg10:*

70: gb_htg11:*

71: gb_htg12:*

72: gb_htg13:*

73: gb_htg14:*

74: gb_htg15:*

75: gb_htg16:*

76: gb_htg17:*

77: gb_htg18:*

78: gb_htg19:*

79: gb_htg20:*

80: gb_htg21:*

81: gb_htg22:*

82: gb_htg23:*

83: gb_htg24:*

84: gb_htg25:*

85: gb_prl:*

86: gb_prl:*

87: gb_prl:*

88: gb_prl:*

89: gb_prl:*

90: gb_prl:*

91: gb_prl:*

92: gb_prl:*

93: gb_prl:*

94: gb_prl:*

95: gb_prl:*

96: gb_prl:*

97: gb_prl:*

98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	745	98.9	1121	5	AF045610
2	742	98.5	1750	6	DME18048
3	592	78.6	1246	5	AF185268
4	592	78.6	50829	65	AC017759
5	592	78.6	232380	5	AE003724
6	538.8	71.6	157133	60	AC007814
7	290.4	38.6	102657	60	AC007891
8	186.8	24.8	747	5	AF226281


```

QY 661 tcggaagacagcatatcgcagccatctcaagtaacacagccctcgctgctgtg 720
Db 785 TCGGAGACAGCATATTCGACCCCATCTCAAGTAAACCAAGCCCTCGCGTCTGGTG 844
QY 721 tcgcccagacagcccgagaggtcgccagataa 753
Db 845 TCGCCGACGACGCCAGAGGGTCCGACAGATTA 877

RESULT 2
LOCUS DME18048 750 bp mRNA INV 03-MAR-2000
DEFINITION Drosophila melanogaster mRNA for deoxyribonuclease, partial.
ACCESSION Y18048
VERSION Y18048.1 GI:4741227
KEYWORDS deoxynuclease kinase; dnk gene.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE Munch-Petersen, B., Knecht, W., Lenz, C., Sondergaard, L. and Piskur, J.
AUTHORS 1 (bases 1 to 750)
TITLE Functional expression of a multistubstrate deoxyribonuclease
kinase from Drosophila melanogaster and its C-terminal deletion
mutants
JOURNAL J. Biol. Chem. 275 (9), 6673-6679 (2000)
MEDLINE 20158988
REFERENCE 2 (bases 1 to 750)
AUTHORS Piskur, J.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1998) J. Piskur, Department of Microbiology,
Technical University of Denmark, DTU - 301, DK - 2800, Lyngby,
DENMARK

FEATURES
source
Location/Qualifiers
1..750
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/dev_stage="embryo"
1..750
/gene="dnk"
1..2750
/gene="dnk"
/codon_start=1
/product="deoxynuclease kinase"
/protein_id="CAB41881.1"
/db_xref="GI:4741228"
/translation="MAEASCARSGTGAETGPTVLIENISGKTYTLNHFEEK
NDICLTPEVKRWNVNGVNLLETKDPKRNAPFQSYVTLMLQSHPTPKLKI
MERSTFSAKCYEVENRNGSLBOGNTLEWKFTLESHVQADLIILAKSPYEA
VERIQKARSESCVPLKTIQELHELDMLIHQRROSKVILVDADLNLNIGTEY
QSESSIDALISSNQPSPLVSPSKRQVAR"

BASE COUNT 188 a 213 c 214 g 135 t
ORIGIN

Query Match 98.5%; Score 742; DB 6; Length 750;
Best Local Similarity 99.3%; Fred. No. 6.7e-158;
Matches 745; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 atggcggagcgacatccctgtcccgaaaggagacagtaacgagggagacccaccc 60
Db 1 ATGGCGGAGCGACATCCGTGTCGCCGAAGGGAGACCAAGTACGCCGAGGCGACCCACCC 60
QY 61 ttacacgtctcatatcgagggcaacatcgagcgaggaaacacagtaattgaacacttc 120
Db 61 TTACACGCTCTCATATCGAGGGCAACATCGGACGCGGAGAACACAGTATTGAACCACTTC 120
QY 121 gagaagtaacaagaacacatctgctgctgtagcagagcccgcgaggaagtggcgcaactc 180
Db 121 GAGAAGTACAAGAAGACATTTGCTGCTGACCAAGCCCGTGGAGAGTGGCGCAACGTC 180

```

```

QY 181 aacggggtlaaatctgtcgtgagctgtagtgcacaagaatcccaagaagtggccatgccctt 240
Db 181 AACGGGGTAAATCTCTGTGAGCTGATGTACAAAGATCCCAAGAGTGGGCGATGCCCTTT 240
QY 241 cagagttatgtacagctgacatgctgcagtcgcacacccgcccacaagaagacta 300
Db 241 CAGAGTTATGTACAGCTGACATGCTGAGTGCAGACACCGCCCAACCAAGAAAGACTA 300
QY 301 aaataaraggagcgctccatttttaagctcgtcctcattgtctcgtlgaagaacatgcagca 360
Db 301 AAATAATAGGAGCGCTCCATTTTTAAAGCTCGCTATGCTTCTGGAGAAACATGCACCA 360
QY 361 aacggctcgtctgagacagggcatgtacaatacgtctgaggagtggtacaaatgcacgaa 420
Db 361 AACGGCTCGTCTGAGACAGGGCATGTACAATACTGTGAGAGAGTGTACAAATTCATCGAA 420
QY 421 gattccattcacctgcgcggcggaacctacatatactgcacacctgcgcggagtgagc 480
Db 421 GATTCATTCACTGCGACGCGGACCTCATATATCTGCCACTGCGCGGAGTGGCGG 480
QY 481 tacgaacgcataccgagcgagcgctcgtctgaggagagcgctgcgcgttaagtacct 540
Db 481 TACGAACGCATCCGCGACGCGGCGCTGTTGAGAGAGAGCTGCGCTTAAATACCTT 540
QY 541 caggagctgcatagtgtgcacacaggaactggtgatatcacacagagacgcagctgtgc 600
Db 541 CAGGAGCTGCATGATGAGTTGCACAGAGACTGTTGATACACAGAGACGCAGCTGTGC 600
QY 601 aagcttcctagctcctcagatgcgcagctcgaacttgaaaacattggcagcgagtaacgcgc 660
Db 601 AAGTCTCTAGTCTCTCGATGCGATCTGAACTGGAACCTGGAATAATTGGACCCAGATACGCGC 660
QY 661 tcggaagacagcatatcgcagccatctcaagtaacacagccctcgcggttgctgtg 720
Db 661 TCGGAGACAGCATATTCGACCCCATCTCAAGTAACCAACACCCCTCGCGGTTCTGGTG 720
QY 721 tcgcccagacagcccgagaggtcgccaga 750
Db 721 TCGCCGACGACGCCAGAGGGTCCGACGA 750

RESULT 3
LOCUS AF185268 1246 bp DNA INV 29-FEB-2000
DEFINITION Drosophila melanogaster deoxyribonuclease kinase gene, complete
cds.
ACCESSION AF185268
VERSION AF185268.1 GI:5924307
KEYWORDS fruit fly.
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE Munch-Petersen, B., Knecht, W., Lenz, C., Sondergaard, L. and Piskur, J.
AUTHORS 1 (bases 1 to 1246)
TITLE Functional expression of a multistubstrate deoxyribonuclease
kinase from Drosophila melanogaster and its C-terminal deletion
mutants
JOURNAL J. Biol. Chem. 275 (9), 6673-6679 (2000)
MEDLINE 20158988
REFERENCE 2 (bases 1 to 1246)
AUTHORS Sondergaard, L., Lenz, C., Piskur, J. and Munch-Petersen, B.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-1999) Institute of Molecular Biology, University
of Copenhagen, O. Farimagsgade 2A, Copenhagen K DK-1353, Denmark
FEATURES
source
Location/Qualifiers
1..1246
/organism="Drosophila melanogaster"
/strain="Oregon R"
/db_xref="taxon:7227"
/chromosome="3R"

```


mRNA
 KYLPSTDDDELTYSGKNAPQDSAKANHLETSTHNSGIIYLLNTEIRSSGGLMSSST
 DPMSKTSMPSCSGVQKRLCKCGOKFLAAL"
 join(37389..37519,37616..38532,38615..38648,38703..>38798)
 /gene="Rh2"
 /product="CT33926"
 /db_xref="FLYBASE:FBan0016740"
 /db_xref="FLYBASE:FBgn0003248"
 <37389..>38798
 /gene="Rh2"
 /note="CG16740"
 /map="91D3-91D4"
 /db_xref="FLYBASE:FBan0016740"
 /db_xref="FLYBASE:FBgn0003248"
 join(37421..37519,37616..38532,38615..38648,38703..38798)
 /gene="Rh2"
 /note="Rh2 gene product"
 /codon_start=1
 /db_xref="FLYBASE:FBan0016740"
 /db_xref="FLYBASE:FBgn0003248"
 /protein_id="AAF55601.1"
 /db_xref="GI:7300445"
 /translation="MERSHLEPEPEDLAHSGPRFOAOSGSGSVLDNVLPMALVNP
 YKSRPAPDPMASKITGLFTLALIMISCCGNGVYVYFGGKSLRTPANLVNLAFLS
 DFCMAASPVMIINFTYETVYLPLMCDIYAGGSGFGCVSISMCMIAFDYNYIV
 KGINGTPMTIKTSLKILFLMMAVFNTPVPLISMYVBPBGNITACSIDYRMNMP
 RSLITYSIFVYTPFLICYSWFIIAVALHAKAREOKKKNVKSLSSEDCRS
 AEGAKVALTITSLMAMPTVLVLCFGLFKIDGLPLPLTWGATFAATSAVYNI
 VYGISPKRYIVLIEKPCMCVGNMDEKPKPADPSDETTSSEADSKA"
 complement(<39713..>40045)
 /gene="CG14294"
 /product="CT33923"
 /db_xref="FLYBASE:FBan0014294"
 /db_xref="FLYBASE:FBgn0038656"
 /evidence=not_experimental
 complement(<39713..>40045)
 /gene="CG14294"
 /map="91D3-91D4"
 /db_xref="FLYBASE:FBan0014294"
 /db_xref="FLYBASE:FBgn0038656"
 /evidence=not_experimental
 complement(39713..40045)
 /gene="CG14294"
 /note="CG14294 gene product"
 /codon_start=1
 /db_xref="FLYBASE:FBan0014294"
 /db_xref="FLYBASE:FBgn0038656"
 /evidence=not_experimental
 /protein_id="AAF55602.1"
 /db_xref="GI:7300446"
 /translation="MAKAVASKNYDDQHRHTASRAKNVDAQIHDFYVLENOKV
 AAREVDSAYETTKCIEVFYSGELEQFKDCIEHVTIRLMERLSLQITTKRQASG
 ASHLKTIW"
 complement(<40400..40729,40945..>40953)
 /gene="CG14293"
 /product="CT33922"
 /db_xref="FLYBASE:FBan0014293"
 /db_xref="FLYBASE:FBgn0038657"
 /evidence=not_experimental
 complement(<40400..>40953)
 /gene="CG14293"
 /map="91D3-91D4"
 /db_xref="FLYBASE:FBan0014293"
 /db_xref="FLYBASE:FBgn0038657"
 /evidence=not_experimental
 complement(join(40400..40729,40945..40953))
 /gene="CG14293"
 /note="CG14293 gene product"
 /codon_start=1
 /db_xref="FLYBASE:FBan0014293"
 /db_xref="FLYBASE:FBgn0038657"
 /evidence=not_experimental
 /protein_id="AAF55603.1"
 /db_xref="GI:7300447"
 CDS

mRNA
 /translation="MEATGSGVGNVAVOOSIDMANEOFKIAQVVAQKLPNVSEARN
 DANDQDITLKLALSHCEAFELKSTQGVLDLHKITCVKAVPFGFTALDRLAHEHPIYGA"
 SGAARIGFPC"
 complement(join(42879..43420,43496..43595))
 /gene="CG14292"
 /product="CT33921"
 /db_xref="FLYBASE:FBan0014292"
 /db_xref="FLYBASE:FBgn0038658"
 complement(<42879..>43595)
 /gene="CG14292"
 /map="91D3-91D4"
 /db_xref="FLYBASE:FBan0014292"
 /db_xref="FLYBASE:FBgn0038658"
 complement(join(42968..43420,43496..43549))
 /gene="CG14292"
 /note="CG14292 gene product"
 /codon_start=1
 /db_xref="FLYBASE:FBan0014292"
 /db_xref="FLYBASE:FBgn0038658"
 /protein_id="AAF55604.1"
 /db_xref="GI:7300448"
 CDS

Query Match 78.6% Score 592; DB 5; Length 232380;
 Best Local Similarity 84.4%; Pred. No. 2,6e-124;
 Matches 748; Conservative 0; Mismatches 5; Indels 133; Gaps 2;

QY 1 atggcgagagcgagcgtccctgctgcccgaaggggaaccaagtaagcgagggcaccgacc 60
 Db 119516 atggcgagagcgagcgtccctgctgcccgaaggggaaccaagtaagcgagggcaccgacc 119575
 QY 61 ttaccggtccatcgagggcaacatcgagcgagggaaagacacgatttgaacacattc 120
 Db 119576 ttaccggtccatcgagggcaacatcgagcgagggaaagacacgatttgaacacattc 119635
 QY 121 gaaataaagaacagacattgctgctgagcgagcgccctcgagaagtgcgcaagtc 180
 Db 119636 gaaataaagaacagacattgctgctgagcgagcgccctcgagaagtgcgcaagtc 119695
 QY 181 aacgggtaaatctgctgtagctgtagtgaacaaagatcccaagaagtggccatcc 240
 Db 119696 aacgggtaaatctgctgtagctgtagtgaacaaagatcccaagaagtggccatcc 119755
 QY 241 cagaagttagtcgctgacatgctgtagcgagtcgacacgcccacaacaagaagta 300
 Db 119756 cagaagttagtcgctgacatgctgtagcgagtcgacacgcccacaacaagaagta 119815
 QY 301 aaataargagcgctccattttagcgctc----- 331
 Db 119816 aaataargagcgctccattttagcgctc----- 119875
 QY 332 -----gctattgctcgt 344
 Db 119876 tgaccgaamaacgcttaaacgaactatccattaccacgcccgcagctgattgcttct 119935
 QY 345 ggaataacatcgagaaagcgctgctgtagcagagcagtcagtaacatcgctggaagagtg 404
 Db 119936 ggaataacatcgagaaagcgctgctgtagcagagcagtcagtaacatcgctggaagagtg 119995
 QY 405 gtacaagttaatcgagaagtgatcaactgagcgagcgagcctcatcatatattcgcgac 464
 Db 119996 gtacaagttaatcgagaagtgatcaactgagcgagcgagcctcatcatatattcgcgac 120055
 QY 465 ctgcgcgagtgtagcagcaacgcatccgagagggcgctgcttgaagagagtcgct 524
 Db 120056 ctgcgcgagtgtagcagcaacgcatccgagagggcgctgcttgaagagagtcgct 120115
 QY 525 gcccttaagtaacctcaagagtgtagtgcagccagagtgtagtaacacagag 584
 Db 120116 gcccttaagtaacctcaagagtgtagtgcagccagagtgtagtaacacagag 120175
 QY 585 acgacgagtcgtgca----- 601
 Db 120176 acgacgagtcgtgca----- 120235

*	3795	3874: gap of unknown length
*	3875	4760: contig of 886 bp in length
*	4761	4840: gap of unknown length
*	5840	5539: contig of 799 bp in length
*	5841	5719: gap of unknown length
*	5720	6629: contig of 910 bp in length
*	6630	6709: gap of unknown length
*	6710	8061: contig of 1352 bp in length
*	8062	8141: gap of unknown length
*	8142	8980: contig of 839 bp in length
*	8981	9060: gap of unknown length
*	9061	10080: gap of unknown length

[illegible]

15723	gap of unknown length
15805	contig of 1046 bp in Length
16851	gap of unknown length
16931	contig of 572 bp in length
17503	gap of unknown length

16631	17502	gap of unknown length
16703	17522	contig of 572 bp in length
17803	18752	gap of unknown length
17883	18793	contig of 1131 bp in length
18114	18793	gap of unknown length
18194	19418	contig of 625 bp in length
19419	19458	gap of unknown length
19499	20606	contig of 1108 bp in length
20607	20686	gap of unknown length
20687	21333	contig of 647 bp in length
21334	21413	gap of unknown length
21414	22148	contig of 735 bp in length
22149	22228	gap of unknown length
22229	23167	contig of 828 bp in length

21334	21413	gap of unknown length
21414	22148	contig of 735 bp in length
22149	22228	gap of unknown length
22229	23167	contig of 939 bp in length
23168	23247	gap of unknown length
23248	24319	contig of 1072 bp in length
24320	24399	gap of unknown length
24320	24489	contig of 490 bp in length
24490	24890	gap of unknown length
24890	25811	contig of 612 bp in length
25812	25661	gap of unknown length
25662	26596	contig of 935 bp in length
26597	26676	gap of unknown length
26677	27576	contig of 900 bp in length
27577	27856	gap of unknown length
27857	28482	contig of 723 bp in length

26599	26576:	gap of 952 bp in length
26677	27576:	contig of 900 bp in length
27577	2756:	gap of unknown length
27657	28387:	contig of 731 bp in length
28388	28487:	gap of unknown length
28468	29554:	contig of 1087 bp in length
29555	29634:	gap of unknown length
29635	30026:	contig of 992 bp in length
30627	30706:	gap of unknown length
30707	31924:	contig of 1218 bp in length
31925	32004:	gap of unknown length
32005	33463:	contig of 1459 bp in length
33464	33543:	gap of unknown length
33544	34666:	contig of 1123 bp in length
34667	34746:	gap of unknown length
34747	35596:	contig of 850 bp in length
35597	35676:	gap of unknown length
35677	36672:	contig of 996 bp in length
36673	36752:	gap of unknown length
36753	38105:	contig of 1303 bp in length
38056	38033:	gap of unknown length
38136	39467:	contig of 1332 bp in length
39468	39547:	gap of unknown length
39548	40728:	contig of 1182 bp in length

ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 102657)
 Celniker, S.E., Agbayani, A., Arcaine, T.T., Baxter, E., Blazek, R.G.,
 Burenhoft, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
 Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
 Hinkley, A., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Katta, K.,
 Kearney, L., Lee, B., Lewis, S., Li, P., Lomocan, M.A., Mazda, P.,
 Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S.,
 Richards, S., Sethi, H., Svirskas, R.R., Wan, K.H., Webster, D.,
 Woolley, P., Yang, S., Yee, M., Yu, C. and Rubin, G.M.
 Sequencing of Drosophila melanogaster
 Unpublished
 2 (bases 1 to 102657)
 Celniker, S.E., Agbayani, A., Arcaine, T.T., Baxter, E., Blazek, R.G.,
 Burenhoft, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
 Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
 Hoskins, R.A., Houston, K.A., Hummasti, S.R., Katta, K., Kearney, L.,
 Kim, E., Lee, B., Lewis, S., Li, P., Lomocan, M.A., Mazda, P.,
 Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S.,
 Pfeiffer, B., Poon, L., Sequeira, A., Sethi, H., Snir, E.,
 Svirskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zierman, L.L. and
 Rubin, G.M.
 Direct Submission
 Submitted (22-JUN-1999) Drosophila Genome Center, Lawrence Berkeley
 Laboratory, MS 64-121, Berkeley, CA 94720, USA
 On Feb 11, 2000 this sequence version replaced gi:6598757.
 For further information about this sequence, including its location
 and relationship to other sequences, please visit our sequence
 archive Web site (<http://www.fruitfly.org/sequence/>) or send email
 to bugreport@fruitfly.berkeley.edu. All contigs in this submission meet
 the following cutoffs: length >= 200 bases.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 80 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1
 891 970: contig of 890 bp in length
 971 970: gap of unknown length
 1603 1602: contig of 632 bp in length
 1683 2326: gap of unknown length
 2327 2406: contig of 644 bp in length
 2407 3212: contig of 806 bp in length
 3213 3292: gap of unknown length
 3293 4089: contig of 797 bp in length
 4090 4169: gap of unknown length
 4170 4842: contig of 673 bp in length
 4843 4923: gap of unknown length
 4923 5608: contig of 685 bp in length
 5608 6356: gap of unknown length
 6357 7427: contig of 669 bp in length
 7427 7507: gap of unknown length
 7508 8305: contig of 798 bp in length
 8306 8385: gap of unknown length
 8386 9165: contig of 779 bp in length
 9165 9244: gap of unknown length
 9245 10151: contig of 906 bp in length
 10151 10230: gap of unknown length
 10231 10981: contig of 751 bp in length
 10982 11061: gap of unknown length
 11062 11746: contig of 685 bp in length
 11747 11826: gap of unknown length
 11827 12399: contig of 573 bp in length
 12400 12479: gap of unknown length
 12480 13371: contig of 892 bp in length

13372 13451: gap of unknown length
 13452 14048: contig of 597 bp in length
 14049 14128: gap of unknown length
 14129 15337: contig of 1209 bp in length
 15338 15417: gap of unknown length
 15418 15907: contig of 1490 bp in length
 15908 16987: gap of unknown length
 16988 17749: contig of 762 bp in length
 17750 17829: gap of unknown length
 17830 18763: contig of 934 bp in length
 18764 18844: gap of unknown length
 18845 19605: contig of 763 bp in length
 19606 19687: gap of unknown length
 19688 20783: contig of 1097 bp in length
 20784 20864: gap of unknown length
 20865 22020: contig of 1157 bp in length
 22021 22101: gap of unknown length
 22102 23471: contig of 1370 bp in length
 23472 23550: gap of unknown length
 23551 24806: contig of 1256 bp in length
 24807 25528: contig of 642 bp in length
 25529 25608: gap of unknown length
 25609 26310: contig of 702 bp in length
 26311 26390: gap of unknown length
 26391 27155: contig of 765 bp in length
 27156 27235: gap of unknown length
 27236 27909: contig of 674 bp in length
 27910 27989: gap of unknown length
 27990 29585: contig of 1596 bp in length
 29586 29665: gap of unknown length
 29666 30871: contig of 1206 bp in length
 30872 30951: gap of unknown length
 30952 31776: contig of 825 bp in length
 31777 31856: gap of unknown length
 31857 33197: contig of 1341 bp in length
 33198 33277: gap of unknown length
 33278 34449: contig of 1172 bp in length
 34450 34529: gap of unknown length
 34530 35270: contig of 741 bp in length
 35271 35350: gap of unknown length
 35351 36805: contig of 1455 bp in length
 36806 36885: gap of unknown length
 36886 38097: contig of 1212 bp in length
 38098 38177: gap of unknown length
 38178 39390: contig of 1213 bp in length
 39391 39470: gap of unknown length
 39471 41255: contig of 1785 bp in length
 41256 41336: gap of unknown length
 41337 42858: contig of 1523 bp in length
 42859 42938: gap of unknown length
 42939 44814: contig of 1876 bp in length
 44815 44894: gap of unknown length
 44895 46297: contig of 1403 bp in length
 46298 46377: gap of unknown length
 46378 47921: contig of 1544 bp in length
 47922 48001: gap of unknown length
 48002 50278: contig of 2277 bp in length
 50279 50358: gap of unknown length
 50359 53204: contig of 2846 bp in length
 53205 53284: gap of unknown length
 53285 55654: contig of 2370 bp in length
 55655 57344: gap of unknown length
 57345 57502: contig of 1768 bp in length
 57503 57582: gap of unknown length
 57583 59074: contig of 1492 bp in length
 59075 59154: gap of unknown length
 59155 61499: contig of 2345 bp in length
 61500 61579: gap of unknown length
 61580 63316: contig of 1737 bp in length
 63317 63396: gap of unknown length
 63397 65599: contig of 2203 bp in length
 65600 65679: gap of unknown length

```

* 65680 67509: contig of 1830 bp in length
* 67510 67589: gap of unknown length
* 67590 69271: contig of 1682 bp in length
* 69272 69351: gap of unknown length
* 69352 71396: contig of 2045 bp in length
* 71397 71476: gap of unknown length
* 71477 73961: contig of 2485 bp in length
* 73962 74041: gap of unknown length
* 74042 77825: contig of 3783 bp in length
* 77826 77905: gap of unknown length
* 77906 81981: contig of 4076 bp in length
* 81982 82060: gap of unknown length
* 82061 87894: contig of 5834 bp in length
* 87895 87974: gap of unknown length
* 87975 88554: contig of 580 bp in length
* 88555 88635: gap of unknown length
* 88636 89263: contig of 629 bp in length
* 89264 89343: gap of unknown length
* 89344 90143: contig of 800 bp in length
* 90144 90223: gap of unknown length
* 90224 90844: contig of 621 bp in length
* 90845 90924: gap of unknown length
* 90925 91384: contig of 460 bp in length
* 91385 91464: gap of unknown length
* 91465 92059: contig of 555 bp in length
* 92060 92139: gap of unknown length
* 92140 92798: contig of 659 bp in length
* 92799 92878: gap of unknown length
* 92879 93422: contig of 464 bp in length
* 93433 93422: gap of unknown length
* 93423 94181: contig of 759 bp in length
* 94182 94261: gap of unknown length
* 94262 94816: contig of 555 bp in length
* 94817 94896: gap of unknown length
* 94897 95527: contig of 631 bp in length
* 95528 95607: gap of unknown length
* 95608 96340: contig of 733 bp in length
* 96341 96420: gap of unknown length
* 96421 96851: contig of 431 bp in length
* 96852 96931: gap of unknown length
* 96932 97619: contig of 688 bp in length
* 97620 97699: gap of unknown length
* 97700 98268: contig of 569 bp in length
* 98269 98348: gap of unknown length
* 98349 98998: contig of 650 bp in length
* 98999 99078: gap of unknown length

```

Query Match 38.6%; Score 290.4; DB 60; Length 102657;
 Best Local Similarity 99.7%; Pred. No. 5.3e-56;
 Matches 291; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 atggcgagagcagatcctctgcccgaaggaaggaacgaatcgcgaagagcagccagccc 60
DB 46006 ATGGCGAGAGCAGATCCTCTGCCCCGAAGGAGGACCAAGTACCCGAGGGACCCAGCC 46065
OY 61 ttcaacgctccatcagaggaacatcgcagcgcgggaagacacgatttgaaccactc 120
DB 46066 TTCAACGCTCCATCAGAGGGAACATCGGACGCGGAACCAAGTATTGAACCACTTC 46125
OY 121 gagaagtaacaagacattcctctcgaacgagccgctgagaagatgagcgcaagctc 180
DB 46126 GAGAAGTAAACAAGACATTTGCTCTGCTGACCGCCCTCGAGAGTGGCCCAACGTC 46185
OY 181 aacggggaatctgctgagcgtatgtataaagaatcccaagaagtggccatgccttt 240
DB 46186 AACGGGGAATCTGCTGAGCTGATGTATACAAATAATCCCAAGAAGTGTGCCATGCCCTT 46245
OY 241 cagaattatgtaacgtatgacatgctgcaatgcgacacacgcccacaca 292
DB 46246 CAGAATTATGTACCGCTGACCATGTGTCACACACCGCCCAACACAACA 46297

```

RESULT 8

```

AF226281
LOCUS AF226281 747 bp mRNA INV 02-APR-2001
DEFINITION Bombyx mori putative deoxynucleoside kinase mRNA, complete cds.
ACCESSION AF226281
VERSION AF226281.1 GI:13506750
KEYWORDS
SOURCE
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
Ditrysia; Bombycoidea; Bombycidae; Bombyx.
REFERENCE
1 (bases 1 to 747)
AUTHORS Knecht, W., Piskur, J., and Munch-Petersen, B.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-2000) Department of Microbiology, Technical
University of Denmark, Building 301, Lyngby, DK DK-2800, Denmark
FEATURES
source
location/Qualifiers
1..747
/organism="Bombyx mori"
/db_xref="taxon:7091"
1..747
CDS
1..747
/note="similar to Drosophila melanogaster deoxynucleoside
kinase encoded by GenBank Accession Number Y18048"
/codon_start=1
/product="putative deoxynucleoside kinase"
/protein_id="AAK28318.1"
/db_xref="GI:13506751"
/translation="MSANNVPEFVEGNIGSGKTTLEHFRPFEDITLLEPEVM
EDIKCNLELMYKMDPEKMAVPOSYSILMDHNRPAFPPVRLKERSLFSAKCFV
EIHNNNTLNHROARVLDPEPRFTQHNIDIDLIVLKPSIVORIKRRARSEEO
CVPLSYTELRLHEDWILINRTIAECAPVLYDADLDLSQITDEXKRSHQILRAV
NVVSSPKRHSKPKPISTPIKTPHMKIL"
BASE COUNT 245 a 152 c 155 g 195 t
ORIGIN
Query Match 24.8%; Score 186.8; DB 5; Length 747;
Best Local Similarity 58.8%; Pred. No. 2.9e-32;
Matches 361; Conservative 0; Mismatches 247; Indels 6; Gaps 2;
OY 56 agccctcaccgctcccatcagcaggaacatcgcagcgcgggaagacacgatttgaacc 115
DB 20 AACCATTCACCGTGTCTGTCGAGGATACATAGTACGCGGTAAACACATTTCTGCAAC 79
OY 116 actcgcgaagtaacaagacacattcctctcgaacgagccgctgagaagtggcgca 175
DB 80 ATTTTCGTCACTTTGAGGATATCATCTT--TGTTGACGAGGCCCTTGAAATGTGGCAG 136
OY 176 acgtcaacgggtaaatctctcgtgagcgtatgtataaagatcccaagaagtggccatgc 235
DB 137 ATCTTAAGGTTCATATCTTTTGGAACTCATGTACAAATATCCAGAAATGGCGATGA 196
OY 236 ccttcagaattatgtacagctgacatcgtcagtcgacgcacacgcgcccaacaacaaga 295
DB 197 CATTCACGTCATACGTTTCTTACGATGTTGAGCATGTGACATCGGAGACCTGCTTCAACTC 256
OY 296 agctcaaaaataaggaacgcttcattttagcgctgcgcatctgctcgtgggaacatgc 355
DB 257 CAGTAAAGCTTAATGAGACGATCATTTATTCAGTCCGAGATATCTCTTCTTGAAACATTA 316
OY 356 gacgaagaagcgtcgtgagcagggcagcgtatgtacaaatcgtggaagagtgtaaaattca 415
DB 317 TGAGAAATATATACCTTCATCCAGCAGACAGTTTTCAGTACTGTATGATGATCCCATTTCA 376
OY 416 tcgaagagtcattcattcattcagcagcgaactcatatatactctgcgacctgcgcggag 475
DB 377 TCCAACACACATTCCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 436
OY 476 tggcgttaacgaatcctcgcagcgcgagcgtctgtctttaggaagagcgtcgtgcgcctaa 535
DB 437 TAGGTACCAAGGATTAAGAAAGAGAGCTGTTTACGAAGAGCAGTGTGTGCGCCCTTAT 496
OY 536 acctcaggaagctgcatgagtggtgacccaagagctggttatacaaccaagagcagcgagt 595

```


Db	302	GGTCATTTCACAGCGCAGCAAGATATATTTTTGTAGAAACCTGTATAGAAAGTGGAGAAATCC	361
Oy	374	agcaaggagcattgtaacaatcgcctgtaggagatggtacagttcatctgaagaatccattacc	433
Db	362	CAGAAGTGGACTATGTACTTGTTCGGAAATGTTTGACTGATCTTGAGCAACATGGACG	421
Oy	434	tgcaggcggaacctatcatataltctgcacactcgcgcgaggttgctgtlaagaaacgtacc	493
Db	422	TGTCTGTGATTATATATGTTTAACTTGGACCAATCCTGAGACTGTGTACCAAGGTTAA	481
Oy	494	ggcagcgggctcgtctctgagagagatcgctgtgccgttaagtaccttcaagagatcatg	553
Db	482	AGAAAGAGATCGACGGAAGAGAGAGAAAGTCATTCGGCTGGAAATACCTGGAAGCAATTCACC	541
Oy	554	agttgcaccagactgttgat	575
Db	542	ATTCATCGATGAGAGTGGCTCAT	563

RESULT	12			
AB046005				
LOCUS		3060 bp	mRNA	18-JUL-2000
DEFINITION		Macaca fascicularis brain cDNA, clone:QCCF-13136.		
ACCESSION		AB046005		
VERSION		AB046005.1		
KEYWORDS		GI:9280065		
SOURCE		Macaca fascicularis adult cDNA to mRNA, clone 13h-macmne brain		

ORGANISM

REFERENCE

TITLE	Author
Isolation of full-length cDNA clones from macaque brain cDNA	Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.

JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 3066)
AUTHORS Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-2000) to the DDBJ/EMBL/GenBank databases.

COMMENT

FEATURES

```

source
1. 3060 .....
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone_qc="QCCE-13136"
/clone_lib="macaque brain cDNA library QCCE
/dev_stage="adult"
472. 1269
CDS

```

```

/codon_start=1
/product="unamed protein product"
/protein_id="BAB01587.1"
/db_xref="GI:9280066"
/translation="MLRLPLRGMAALALRCPEPGSPSPASGCSRRVQACMPSPDKE
REKRSKSVICVEGNIASGKTCLEFENADILEYTVSKMRVNRKHNPGLMYDQA
SRNGLITQIVVQMLDMDRHCPQVSSRLRERSTHSARYLVENLVRSQGMPEVDYV
LSPEFQWILDMQVSDIDLIYLRINPTCTQRLKRREREKVIPLYEALIHNLHEE
WLKGSLEFPAALPLVLEADHMRIMQTLCEQNRDRILTPENRKLGP"
BASE COUNT
ORIGIN      721 a      758 c      867 g      714 t

```

Query Match	15.1%	Score 113.4;	DB 85;	Length 3060;
Best Local Similarity	52.4%	Pred. No. 9.8e-16;		
Matches 299; Conservative	0;	Mismatches 266;	Indels 6;	Gaps 2;

```

0y      74   tcgaggcaacatctcggcagcgggaagaaccagtatttgaaccacttcgagaagtacaaga    133
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      635   TCGAGGGCATAATTGCAAGTGGGAAGCAGCATGCCCTGGAATTTCTTCCA--ATGCGA    691

```

Db 692 CAGACATTGAGGTAATTACGCGAGCCCTGTCCAAAGTGGAGAAATGTCCTGGCCACATC 751

D_b

752 CTTGGGCGCTGATGTACCAAGATGCCCTCGCTGGGGCCTTACGCCTAACAGACTTACGTGC 811

Db 812 AGCTCACCATGCTGGACAGGCATACCTGTCTCAGGTGTCATCTGTACGTTGATGAGACA 871

Oy 314 gctccatctttagcgcctcgcctatctgttcgtgaagaacatagcagcaaacctcctctc 373

D_b 872 GGTGATTACACCGCAGATATACATTTTGTAGAAACCTGTATAGAGTGGGAAGATGC 931
O_y 374 agcaggcattgtacataacgcgtggaagagtggtacaagttcatctgaagagtcattcac 433

434 **tg**cagcgcgaccctcatcatatatactctgcgacactgcgcgaagtgycgtacgaacgatcc 493
 932 **CAG**AAGTGGACTATAGTACTCTCTGTGGGAATGGTTTGACTGGATCTTTGAGGAACATGGACG 991

[illegible][illegible]

Qy	611	tcctcgatgcgatctgaacctggaacat	641
Db	1172	TGATCGAGGCTGACCAACCATGAGAGGAT	1202

RESULT 13	
HSU77088	
LOCUS	
HSU77088	1060 bp
	5000
	5000

DESCRIPTION	human clym4laine kinase 2 (TK2) mRNA, complete cds.
ACCESSION	U77088
VERSION	U77088.1
KEYWORDS	GI:1905968

ORGANISM	REFERENCE
Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostom	
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
1 / base 1 to 1000	

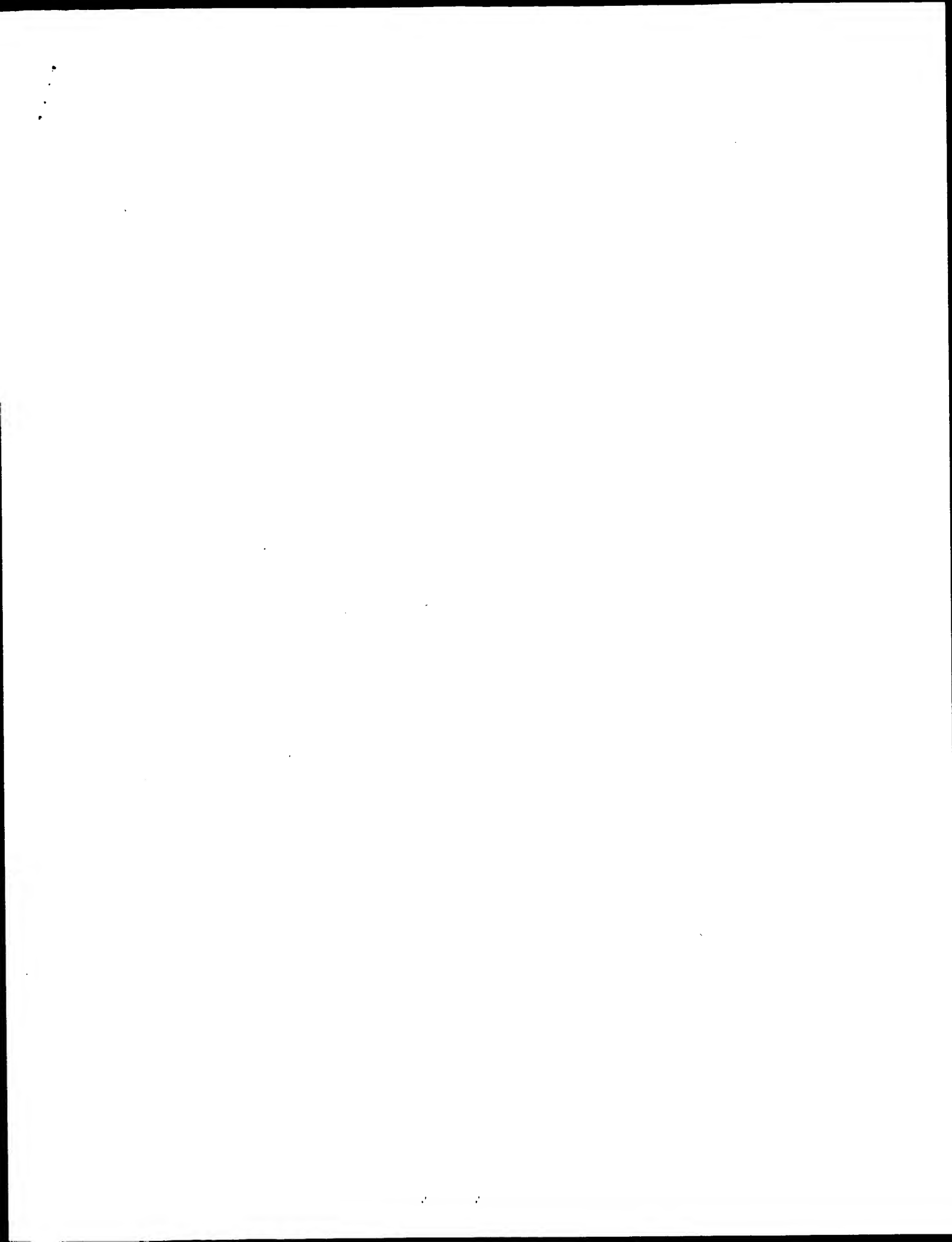
FOUNDRY TITLE	FOUNDRY TITLE
Jonasson, M. and Karlsson, A. Cloning of the CDNA and chromosome localization of the gene for human thymidine kinase 2	

JOURNAL	J Biol. Chem.	272 (13),	8454-8458	(1997)
MEDLINE	97236800			
REFERENCE	2 (bases 1 to 1960)			
AUTHORS	Johansson,M. and Karlsson,A.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-NOV-1996) Medical Biochemistry and Biophysics,			
	Karolinska Institute, Doktorsringen 2A, Stockholm 171 77, Sweden			
FEATURES	Location/Qualifiers			
SOURCE	1..1960			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="4"			
gene	9..713			
	/gene="TRK2"			
CDS	9..713			
	/gene="TRK2"			
	/function="deoxyribonucleoside kinase"			
	/codon_start=1			
	/product="Thymidine kinase 2"			
	/protein_id="AAC51167.1"			
	/db_xref="gi:1905969"			
	/translation="MGAFICOPSSDKEEKRSVLCEGNIAGKTTCLEFFSNATDVEVIFEPYSKRWNRGNPLGIAMDHSRMLTLOTYVOLMDLRHPQVSRYLMEHSISARITPEINLYRSGKMPEVDYVSLISEFDMLIRMDVSVLIYLRTNPETCYGRFLKRCKREKVIVLELTLEAIHHHEEULINGSLFPAAPLVYLEADHHMERMLELEONRRILLTPENRKHCIP"			
BASE COUNT	471 a	465 c	525 g	499 t
ORIGIN				
Query Match	15.0%;	Score 113.2;	DB 97;	Length 1960;
Best Local Similarity	53.0%;	Pred. No. 1.2e-15;		
Matches 266; Conservative	0;	Mismatches 233;	Indels 3;	Gaps 1;
Oy	74	tccgaagcaacatcgcgcaacggagaacacgactatttgacaaccattcgagaagtacaaaga	133	
Db	79	tccgaggccaattttgcaggtgggaagacacacatgccctggatttcttcacca--acggca	135	
Oy	134	acgaacattgcctcgtcgtaccgcgagcccgttcgagaagtgcgcgaacgtcaacggyttaaac	193	
Db	136	cagacctgcagagtggttaacggagacctgtgtccaaagtgagaataatgtccctggccacacattc	195	
Oy	194	tgcctgaagctgatgtacaaagatccccaaagaatgggcattgccccttcagagattatgtca	253	
Db	196	ctctggggcctgaatgacacacagatgccttcgcgtggggctgttacgctaacagacttatgtgc	255	
Oy	254	cgcctgaccatcgtcgtcagtcgcacacccgcccaaccaaagaagctaataatargagc	313	
Db	256	agcttcacacatgcgtgcagcagcacatgacgcgtccagctgacatgtagcaggttatggaaa	315	
Oy	314	gtcccaatttttagcgttcgtctatgttcctgttgagacatatgcgaagaacgcgtcgtcg	373	
Db	316	ggctcgatttcacacgcgacagattacatttttctagaaaaacctgtagaagtggaabattgc	375	
Oy	374	agcaaggaacgtlaacaatacagcttggaagatgtygtacaagtcatcogaagaagtccatcacc	433	
Db	376	cacaaagtgacatmgatagttctgtcggaatggtttgactgtagactcttgaggaacatgatgacg	435	
Oy	434	tgcgaagcggaacctcattcataatctatctgcgcacctccgcggaaggtggcgctlaagaacgatcc	493	
Db	436	tgtctgttgtgattgatagttttaccttcgcgacccaatctctgagacttgatttaccagaaggttaa	495	
Oy	494	ggcagcgcggtcgtcttcaggaagacgtgcgttcgcgttlaagtaaccttcaggaagctgcaty	553	
Db	496	agaaagatgcacaggaagagaagaaagtcattccgctggaaatnactcggaaagcaatttcacc	555	
Oy	554	aattgcacacgaactggtgat	575	
Db	556	atctccatgagagtggtcat	577	

LOCUS	HSU80628	2710 bp	mRNA	PRI	24-MAR-1997
DEFINITION	Human thymidine kinase 2 isoform B (TK2) mRNA, alternatively spliced, partial cds.				
ACCESSION	U80628				
VERSION	U80628.1	GI:1905973			
KEYWORDS	.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 2710)				
AUTHORS	Johansson,M. and Karlsson,A.				
TITLE	Cloning of the cDNA and chromosome localization of the gene for human thymidine kinase 2				
JOURNAL	J. Biol. Chem. 272 (13), 8454-8458 (1997)				
MEDLINE	97236800				
REFERENCE	2 (bases 1 to 2710)				
AUTHORS	Johansson,M. and Karlsson,A.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-DEC-1996) Medical Biochemistry and Biophysics,				
	Karolinska Institute, Doktorarringen 2A, Stockholm 171 77, Sweden				
FEATURES	Location/Qualifiers				
source	1..2710				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="4"				
	<790..1464				
gene	/gene="TK2"				
	<790..1464				
CDS	/gene="TK2"				
	/function="deoxyribonucleoside kinase"				
	/note="Similar to 3' region of TK2 of GenBank Accession				
	Number U77088; defectively spliced mRNA that is probably				
	not translated"				
	/codon_start=1				
	/product="thymidine kinase 2 isoform B"				
	/protein_id="AAC51168.1"				
	/db_xref="GI:1905974"				
	/translation="DKDEKREKSVICVEGNIAGKTTCLEFFSNATIDVLEVLPVK				
	WRNVGHNPDLIMHDNRKMTITLOTYOLMLDRHPPOVSYSRMSRSIHSARYLFI				
	VENLRSGKEPEVDIVLVSEFMDILRMADYVSVLIYLRNPEPCYORLKRCREEE				
	KVIPLLEYLAIHHEWLIKGSLEFPMAAPVLIEADHNEMLELFQONDRILTFE				
	NRKHP"				
BASE COUNT	634 a 666 c 754 g 656 t				
ORIGIN					
Query Match	15.0%; Score 113.2; DB 97; Length 2710;				
Best Local Similarity	53.0%; Pred. No. 11e-15;				
Matches 266; Conservative	0; Mismatches 233; Indels 3; Gaps 1;				
QY	74 tcgaaggcaacatcgcgcaacgggaagaccagcatatttgacaacctcgagaagtacaaga 133				
Db	830 TCGAGGGCATATTTCGACGTGGGAACAACGACATGCCTGGAAATTTCTTCCCA--ACGCGA 886				
QY	134 acgacatlctgcctcgtgtaaccgagcccgctcgagaagtggcgcaacgtccaacgggtaaac 193				
Db	887 CAGAAGCTCGAGTGTTAACGAGACCTGTGTGCCAAGTGGGAATAATGTCGTGGCCACAAATC 946				
QY	194 tgttgaagctcgatgatacaaatgcccaagaagtyggcgcaatgccccttcagagttatgtla 253				
Db	947 CTCGGGCTCGATGTACACAGATGCTCTCGCTGGGGGTCTTAACGTCACAGACTATANGTGC 1006				
QY	254 cgctgaccatctgctcgauctgcaacacgcgcccaacaocaagaagaagctaataargagc 313				
Db	1007 AGCTCACACATCGTGGACAGACGATACGTCGTCAGTGTGCATCTGACGTTGATGTGAAGA 1066				
QY	314 gtctcattttaagcgctcgcgatattgtctctgtggaagacatcgcaagcaaaagcctcgctg 373				
Db	1067 GGTCGATTTCACAGGGCAGATATCACTTTTTTACAAAAACCTGTTAGAAAGTGGGAANAATGC 1126				
QY	374 agcagggatcatgataacatcacgctggaagagtgatgataaagttcatacogaagaagtcattcac 433				

OY		310	gagcgcctccatctttttagcgcgctcgtattgcttcgtlygaacaacatcygacgaacgctcg	365
Db		426	GAGAGGCTGTGTACAGTGACAGGTATATTCTTTCGAAGAATCTATTGAAATGGCTCC	485

Search completed: October 28, 2001, 06:25:09
Job time: 6665 sec




```

32 acilysserthrphenylalanylglutylglutylslylasprrpg 49
143 GCCTGCTGACCGCCGCGAGAGTGGCGAACGTC..... 180
49 luValValProGluProValAlaIaArgTrpCysAsnValGlnSerThrGln 65
181 .....AACGGGTAAATCTGCT 197
66 ASPGluPheGluGluLeuThrMetSerGlnLysAsnGlyGlyAsnValle 82
198 GGAGCTGATGTACAAAGATCCCAAGAGTGGCGCATGGCTTTCAGAGTT 247
82 uGlnMetMetTrpGluLysProGluArgTrpSerPheThrPheGlnThrT 99
248 ATGTCAAGCTGACCATGTGTCAGTGGCACCGCCCAACCAACAAGAG 297
99 yAlaLysLeuSerArgIleArgAlaGlnLeuAlaSerLeuAsnGlyLys 115
298 CTAAGA.....ATARGAGCGCTCCATTCTTAG 326
116 LeuLysAspAlaGluLysProValLeuPhePheGluArgSerValLysSe 132
327 CGCTGCTATTGCTGCTGGAGAACATGCGAAGAACGCTGCTGGAGAC 376
132 rAspArgTrpIlePheAlaSerAsnLeuTrpGluSerGluCysMetAsnG 149
377 AGGGCATGTACAAATACGCTGGAGAGTGGACAACTTCATCGAAGAGTCC 426
149 IuThrGluTrpTrpIleTrpGlnAspTrpHisAspTrpMetAsnGln 165
427 ATT.....CACTGACGGGCGACTATCATATATCTGGCACTCGCC 470
166 PheGlyGlnSerLeuGluLeuAspGlyIleIleTrpLeuGlnAlaThrPr 182
471 GGAGTGGCTGACGAGACGATCCGGGAGGCTGCTGACGAGAGT 520
182 oGluThrCysLeuHisArgIleTrpLeuArgGlyArgAsnGluGlnG 199
521 GCGTGCCTTAAGTACCTTCAGAGCTGCATGATTCGACCAAGAGCTGG 570
199 IyIleProLeuGluTrpLeuGluLysLeuHisTrpLysHisGlnSerTrp 215
571 TTGATACACCAAGACGACG.....CAGTCGTGGCA 602
216 LeuLeuHisArgThrLeuLysThrAsnPheAspTrpLeuGlnGluValPr 232
603 GGCTCTAGTCCGATGGCGGATGTAACCTGAGAAACATTCGGCAGAGT 652
232 oLleuThrLeuAspValAsn.....GluAspPheLysAspLysT 246
653 ACCAGCGCTCGAGAGACGATATTTCAGCGCATCTCAAGT 693
246 yGluSerLeuValGluLysValLysGluPheLeuSerThr 259
seq_name: /cgr2_6/ptodata/2/iaa/5B_COMB.pep:US-08-879-561-7
seq_documentation_block:
: Sequence 7, Application US/08879561
: Patent No. 5817482
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Hawkins, Phillip R.
: APPLICANT: Guejler, Karl J.
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESS: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA

```

```

: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: Fastseq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/879,561
: FILING DATE: Herewith
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0325 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 260 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: Genbank
: CLONE: 1480198
: US-08-879-561-7

alignment_scores:
Quality: 333.00 Length: 233
Ratio: 2.297 Gaps: 5
Percent Similarity: 62.232 Percent Identity: 32.618

alignment_block:
US-09-416-579A-1 x US-08-879-561-7 ..
Align seg 1/1 to: US-08-879-561-7 from: 1 to: 260

49 GGCACCCAGCCCTTCACCGCTCATTCAGAGGCGACATCGGCGAG 98
111 111 111 111 111 111 111 111 111 111 111 111
18 GLyArgGlyProArgArgLeuSerIleGluGlyAsnIleAlaValGly 34
99 GACCAAGTATTGTAACCACTTCGAGAGTACAAAGACATTCGCTG 148
111 111 111 111 111 111 111 111 111 111 111
34 sSerThrPheValLysLeuLeuThrLysThrTrpProGluTrpHisVal 51
149 TGACCGAGCCCGTCGAGAGTGGCGCAACGTCAAGGGGTA..... 189
111 111 111 111 111 111 111 111 111 111 111
51 IaThrGluProValAlaIaThrTrpGlnAsnIleGlnAlaIaGlyThrGln 67
190 .....AATCTGCTGGAGCTGATGTACAA 212
68 IyAlaCysThrAlaGlnSerLeuGlyAsnLeuLeuAspMetLeuTrpAla 84
213 AGATCCCAAGAGTGGCGCATGCCCTTCAGAGTATATCATCGCTGACA 262
84 gGluProAlaArgTrpSerTrpThrPheGlnThrPheSerPheLeuSer 101
263 TCCTGCAAGTGCACACCGCCCAACCAACAAGAGCTA..... 300
111 111 111 111 111 111 111 111 111 111 111
101 rGluLeuLysValGlnLeuGluLysProPheProGluLysLeuGlnAlaArg 117
301 .....AAATAARGAGCGCTTCATTTTACGCGCTGCTTCTT 341
118 IyProValGlnIlePheGluArgSerValTrpSerAspArgTrpIlePhe 134
342 CGTGGAGACATGCGACGAAGAGGCTCGCTGGAGCAGGCGATGTACAATA 391

```


[illegible]

Align seg 1/1 to: US-08-860-995-2 from: 1 to: 172

```

49 GGCACCCAGCCCTTCACCGTCCATCGAGGCGCAACATCGCCGCGGAA 98
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
18 G1yArgG1yProArgArgLeuSerIleGluG1yAsnIleAlaValG1yLy 34
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
99 GACCAAGTATTTGAACACCTTCGAGAGTACAAGACGACATTCGCTGC 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
34 SserThrPheValIyLeuLeuThrIySthIyTyrProG1utPrhIstValA 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
149 TGACCCAGCCGTCGAGAGTGGCGGACGTCACCGGGA..... 189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 IatHrG1uProValAlatHrTrG1nAsnIleG1nAlaIaG1yThrG1n 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
190 .....AATGCTGTCGAGTGTGATGACAA 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 LysAlaCysThrAlaGlnSerIeuG1yAsnIleLeuAspMetIyTyr 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
213 AGATCCCAAGAGTGGCCATGCGCTTCAGAGTATATGTCAGCGTGAACA 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
84 G1uProAlaArgTyrSerTyrThrPheGlnThrPheSerPheLeuSerA 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
263 TGCTGAGTCGACACCGCCGCAACCAACAGAGCTA..... 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
101 rG1euLySValGlnIeuG1uProPheProG1uLySLeuGlnAlaArg 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 .....AAATTAARGAGCGCTCCATTTTACGCGCTCGCTATTCGCTT 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
118 LysProValGlnIlePheGlnIuArgSerValIyTyrSerAspArgLeuHsPh 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
342 C 342
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
134 e 134

```

seq_name: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:PCT-US95-00532A-2

```

seq_documentation_block:
; Sequence 2, Application PC/RTUS9500532A
; GENERAL INFORMATION:
; APPLICANT: WEI, ET AL.
; TITLE OF INVENTION: Human Deoxycytidine Kinase 2
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/00532A
; FILING DATE: 13 JAN 95
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-246
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 172 AMINO ACIDS

```

```

; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
PCT-US95-00532A-2

```

```

alignment_scores:
  Quality: 164.50
  Ratio: 2.455
  Percent Similarity: 57.265
  Percent Identity: 32.479

```

alignment_block:

US-09-416-579a-1 x PCT-US95-00532A-2 ..

Align seg 1/1 to: PCT-US95-00532A-2 from: 1 to: 172

```

49 GGCACCCAGCCCTTCACCGTCCATCGAGGCGCAACATCGCCGCGGAA 98
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
18 G1yArgG1yProArgArgLeuSerIleGluG1yAsnIleAlaValG1yLy 34
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
99 GACCAAGTATTTGAACACCTTCGAGAGTACAAGACGACATTCGCTGC 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
34 SserThrPheValIyLeuLeuThrIySthIyTyrProG1utPrhIstValA 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
149 TGACCCAGCCGTCGAGAGTGGCGGACGTCACCGGGA..... 189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 IatHrG1uProValAlatHrTrG1nAsnIleG1nAlaIaG1yThrG1n 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
190 .....AATGCTGTCGAGTGTGATGACAA 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 LysAlaCysThrAlaGlnSerIeuG1yAsnIleLeuAspMetIyTyr 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
213 AGATCCCAAGAGTGGCCATGCGCTTCAGAGTATATGTCAGCGTGAACA 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
84 G1uProAlaArgTyrSerTyrThrPheGlnThrPheSerPheLeuSerA 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
263 TGCTGAGTCGACACCGCCGCAACCAACAGAGCTA..... 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
101 rG1euLySValGlnIeuG1uProPheProG1uLySLeuGlnAlaArg 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 .....AAATTAARGAGCGCTCCATTTTACGCGCTCGCTATTCGCTT 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
118 LysProValGlnIlePheGlnIuArgSerValIyTyrSerAspArgLeuHsPh 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
342 C 342
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
134 e 134

```

seq_name: /cgn2_6/prodata/2/1aa/5B_COMB.pep:US-09-156-836B-2

```

seq_documentation_block:
; Sequence 2, Application US/09156836B
; Patent No. 6242585
; GENERAL INFORMATION:
; APPLICANT: Srivastava, Ranjana
; APPLICANT: Kumar, Deepak
; APPLICANT: Srivastava, Brahm Shanker
; TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS SPECIFIC DNA FRAGMENT
; FILE REFERENCE: U 011876-4
; CURRENT APPLICATION NUMBER: US/09/156, 836B
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 08/997,897
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (4)

```



```

263 ysphehiscysargtyr***sercysargargcysargcyspro..... 277
184 cgttagccttgccgaccttctgcagggctcgccagcagc..... 143
278 .....argcyspnehlcysargglycysargcysglnllycys***se 292
142 .....AATGTCGTTCTTCTACTCTCGAAGTGTTCAATA 106
292 thlsarghargphearfcysargglucyscysan..... Cysargc 306
105 cgtgcttcccgctgcccagatgtgcccctgcagatgagcagcgtaagcgct 56
306 ystpr.....argcysargglucysserarg.....Argpro 316
55 GGGTCCCTCGCGCTACTGTCCTTCGGGACACAGATGCTGCT 8
317 glyleuprocglyargasp***argprovalglyhlsarglysllepro 332
seq_name: /cgn2_6/plodata/2/1aa/6A.COMB.pep:US-08-997-897-2

```

```

seq_documentation_block:
: Sequence 2, Application US/08997897C
: Patent No. 6114514
: GENERAL INFORMATION:
: APPLICANT: SRIVASTAVA, RANJANA
: APPLICANT: KUMAR, DEEPAK
: APPLICANT: SRIVASTAVA, BRAHM SHANKER
: TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS SPECIFIC DNA FRAGMENT
: FILE REFERENCE: u011469-7
: CURRENT APPLICATION NUMBER: US/08/997,897C
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 430
: TYPE: PRT
: ORGANISM: Mycobacterium tuberculosis
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (4)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (6)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (20)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (29)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (54)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (64)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (69)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (89)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (99)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (114)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (119)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (129)

```

```

: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (159)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (169)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (182)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (185)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (219)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (259)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (269)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (291)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (323)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (339)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (349)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (356)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (366)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (410)
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (421)
: US-08-997-897-2

```

```

alignment_scores:
Quality: 141.50 Length: 300
Ratio: 1.252 Gaps: 19
Percent Similarity: 37.667 Percent Identity: 26.000

```

alignment_block:

US-09-416-579a-1/rev x US-08-997-897-2 ..

Align seg 1/1 to: US-08-997-897-2 from: 1 to: 430

```

745 CGACCTCTGCGCTTGCTGGGACACAGACCGGCGAGGCTTTGG 696
||| ||| |||||
88 Arg***sercysargcysrtpcys..... 95
695 TTACTTGAGATGGCGTCGATATGCTGC.....TTCG 664
||| ||| ||| |||||
96 .TyrArgArg***SerArgArgCysCysSerCysProcysCysArgSerP 112
|| .....TGCATATGT..... 638
112 rocys***SerArgArgArg***GlyCysProcysCysSerCysGlnHls 128
|| .....TTCGAGGTTGAGA.....TCGCAATGAG 613
637 .....:|||||||
129 ***GlyCysArgTyrcysargtyrproglyserargtyrproserSerAr 145

```

CLY: PAID ALCO

1997

```

323 AAAATGAGCGCTCCYTTATTTTACCTCTGTGGTGGCGCGGTG 274
841 .....AlaAlaCysCysLThrGlyAlaCys 850
273 CGACTGACGATGTCAGCGCTGACATACCTGAAAGGCGACCT 224
850 lAcysCysAlaThrCysAla..... 856
223 TCTTGGGATCTTGTATCATCAGCTCCAGAGATTACCCCTTACGTTG 174
857 .....CysCysCysLThrCysAlaThrCys.....ThrCys 866
173 CGGCACTTCGACGCGCTCGCTGACGA.....GGCAATGTC 136
866 scYThrAlaThrAlaAlaGlyAlaAlaThrGlyAlaThrGlyCysCys 883
135 GTTCTGTACTTCTGA..... 119
883 hAlaCysThrThrCysThrCysAlaAlaGlyAlaCysCysThr 899
118 ...AGTGTCAAAATACGTGTCTCCGCTCCGATGTCCTCGATG 72
900 GYAlaGlyCysCysThrGlyAlaGlyCysThrCysCys...Thr... 914
71 AGGACGGTGAAGGGCTGGTCCCTGCTACTGTGCTCCCTTGGGC 22
915 .....Gly 915
21 ACAGATGCTGC 10
916 ThrGlyCysCys 919

seq_name: /cgn2_6/plodata/2/1aa/5B_COMB.pep:us-08-705-660-46
seq_documentation_block:
; Sequence 46, Application US/08705660
; Patent No. 5858683
; GENERAL INFORMATION:
; APPLICANT: KEESSE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-JYE
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,660
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: GREENHALGH, DUNCAN A
; REGISTRATION NUMBER: 38,678
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: peptide
US-08-705-660-46

Alignment_scores:
  Quality: 99.00      Length: 248
  Ratio: 0.756      Gaps: 9
  Percent Similarity: 52.823      Percent Identity: 24.194

alignment_block:
US-09-416-579A-1 x US-08-705-660-46

Align seg 1/1 to: US-08-705-660-46 from: 1 to: 515

42 CGCGAGAGCGACCCACCTTCACCGCTCTCATCGAGGCAACATCGCA 91
71 ArgArgHisGluThrArgLeuValGluIleAspAsnGlySerGlnArgG 87
92 GCGGAAGACACAGTATTGAAACCACTTCGAGAACTACACACGACATT 141
87 uPheGlnSerArgLeuAlaAspAlaLeuGlnGlnLeuArgAlaGlnHis 103
142 TGCGTCTGACCGAGCCGCTCGAAGAGTGGCGCAACGTCAACGGGTAA 191
104 ....GluAspGlnValGlnGlnIleTyrLysGlnLeuGlnLysThrTyr 118
192 TCTGCTGAGAGTATGATCAAGA.....TCCCAAGA 223
119 SerAlaLysLeuAspAsnAlaArgGlnSerAlaGlnArgAsnSerAsn 135
224 AGTGGCCATGCGCTTTCAGACTTA..... 248
135 uValGlyAlaAlaHisGlnGlnGlnGlnSerArgIleArgIleAsp 152
249 .....TGTACGCTGACCATGCTGACGTCGACACCGCCCAACCA 290
152 erLeuSerAlaGlnLeuSerGlnLeuGlnLysGlnLeuAlaLysGln 168
291 CAAGAGCTAAATAARAGAGCGCTCCATTTT.....AGG 328
168 uAlaLysLeuArgAspLeuGlnLysSerLeuAlaArgGlnArgAspHis 185
329 CTGCTATTGCTTCGTGAGACATCGAGCAAGAGCGCTCGTGGACAG 378
185 erArgArgLeuLeuAlaGlnLysGlnArgLumetAlaLumetArgAla 201
379 GGCATGTACAATACCTCGAGAGAGTGTACAAATTCAATGAAGATCAT 428
202 ArgmetGlnGlnGlnLeuAspGlnTyrGlnGlnLeuLeuAspIleLys 218
429 TCACCTGACGCGCGACCTCATATAT..... 456
218 uAlaLeuAspmetGlnIleHisAlaTyrArgLysLeuGlnLysGln 235
457 .....CTGGGCACTTCGCGAGGTGGGTACGAACGATCGGCGAG 498
235 LysGlnArgLeuArgLeuSerProSerProHisSerGlnArgSerArg 251
499 CGGCTCGTTCTGAGAGAGC.....TGCCTGCCGT 530
252 ArgAlaSerSerHisSerSerGlnThrGlnLysLysSerValThrLys 268
531 TAAGTACTCTTACGAGCTGATGATGTCACACAGACGATGTTATACAC 580
268 sLysArgLysLeuGlnSerThrGlnSerArgSerPheSerGlnHisA 285
581 AGAGACGACCGCAGCTGTCGAAGGTCTACTCTCGATGCCGAT..... 624
285 laArg.....ThrSerGlyArgValAlaValGlnLysArgGln 299
625 .....CTGAACCTGGAACATTTGCGACCGAGTACACG 657
300 GlyLysPheValArgLeuArgAsnLysSerAsnGlnAspGln 313

```

```

seq_name: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:US-08-989-045-46
seq_documentation_block:
; Sequence 46, Application US/08989045
; Patent No. 6027905
GENERAL INFORMATION:
APPLICANT: KESSEE, SUSAN
APPLICANT: OBAR, ROBERT
APPLICANT: WU, YING-YIE
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
TITLE OF INVENTION: DETECTION OF CERVICAL CANCER
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESS: Testa, Hurwitz & Thibault
STREET: 125 High St.
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989,045
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: GREENHALGH, DUNCAN A
REGISTRATION NUMBER: 38,678
REFERENCE/DOCKET NUMBER: WTP-023 (8395/27)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7100
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 515 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-989-045-46

alignment_scores:
Quality: 99.00 Length: 248
Ratio: 0.756 Gaps: 9
Percent Similarity: 52.823 Percent Identity: 24.194

alignment_block:
US-09-416-579a-1 x US-08-989-045-46
Align seg 1/1 to: US-08-989-045-46 from: 1 to: 515
42 CGCCGAGGCGACCCAGCCCTTCACCGCTCATCGAGGCGCAATCGCA 91
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
71 ArgArgHisGlnThrArgLeuValGlnIleAspAsnGlyLysGlnArgG1 87
92 GCGGGAAGACAGATTTGAACCACTTCGAGAGACTACAAAGACATTT 141
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
87 upheGlnSerArgLeuAlaAspAlaLeuGlnGlnLeuArgAlaGlnHis 103
142 TGCCTGCTGACCGAGCCCGCTCGAGAGTGGCGCACTCAACGGGCTAA 191
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
104 .....GluAspGlnValGlnGlnIleTyrLysLysGlnLeuGlnLysThrTyr 118
192 TCTGCTGAGCTGATGTACAAAG.....TCCCAAGA 223
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
119 SerAlaLysLeuAspAsnAlaArgGlnSerAlaGluArgAsnSerAsnIle 135
224 AGTGGCCCATGCGCTTCAGAGTTA..... 248

```

```

135 uValGlnAlaIleHisGlnGlnIleGlnSerArgIleArgIleAspS 152
249 .....TGTACGCTGACCATCTGCTGACGTCGACACCGCCCAACCA 290
152 erLeuSerAlaGlnLeuSerGlnLeuGlnIleGlnAlaLysGln 168
291 CAGAGACTAAATAATTAAGAGCCCTCCATTTT.....AGCG 328
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
168 uAlaLysLeuArgAspLeuGlnAspSerLeuAlaArgGlnArgAspHis 185
329 CTGGCTATTGCTTCGAGAAATGACGACGAAAGCGCTCGGAGACG 378
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
185 erArgArgLeuLeuAlaGlnLysGlnArgGlnMetAlaGlnMetArgAla 201
379 GGCATGTACATATGCTGGAGAGTGTACAAAGTTTCATTCGAAGTCCAT 428
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
202 ArgMetGlnGlnGlnLeuAspGlnIleTyrGlnIleLeuLeuAspIleLys 218
429 TCACCTGCAGCGGCGACCTCATCATAT..... 456
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
218 uAlaLeuAspMetGlnIleHisAlaTyrArgLysLeuLeuGlnGlnGln 235
457 .....CTGGCACCCTGCGGAGAGTGGCTACGACGATCGGAG 498
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
235 LysGlnArgLeuArgLeuSerProThrSerGlnArgSerArgGly 251
499 CGGCTGCTGCTGAGAGAC.....TGGTCCGCT 530
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
252 ArgAlaSerHisSerSerGlnThrGlnGlnGlnGlnGlnGlnGlnGln 268
531 TAAGTACCTTCAGAGCTGCATGATGTTGACACGACGCTGTGATACAC 580
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
268 sLysArgLysLeuGlnIleThrGlnSerArgSerPheSerGlnHisA 285
581 AGAGACAGCCGACGTCGACAGGCTCTAGTCTCGATCCGAT..... 624
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
285 laArg.....ThSerGlnArgValAlaValAlaValAlaValAspGln 299
625 .....CTGACCTGGAACATTTGGCAGCGAGTACAG 657
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
300 GlyLysPheValArgLeuArgAsnLysSerAsnGlnAspGln 313

seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-08-483-533-41
seq_documentation_block:
; Sequence 41, Application US/08483533
; Patent No. 6172047
GENERAL INFORMATION:
APPLICANT: Roizman, Bernard
APPLICANT: Chou, Joany
TITLE OF INVENTION: Method for Treating Tumorigenic
TITLE OF INVENTION: Diseases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,533
FILING DATE: 07-MAR-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/419,853
FILING DATE: 11-APR-95

```


PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/861,233
 FILING DATE: 31-MAR-92
 ATTORNEY/AGENT INFORMATION:
 NAME: Zeller, James P.
 REGISTRATION NUMBER: 28,491
 REFERENCE/DOCKET NUMBER: 28097/32742
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 41:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 355 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-483-533-41

alignment_scores:
 Quality: 98.00 Length: 237
 Ratio: 1.210 Gaps: 10
 Percent Similarity: 34.177 Percent Identity: 23.629

alignment_block:

US-09-416-579a-1/rev x US-08-483-533-41 ..

Align seg 1/1 to: US-08-483-533-41 from: 1 to: 355

```

622 CGGCATCGAGACTAGACCTTCAGACACTGGCGTCTGTGGTATC 573
    |||||||
6   ArgHisArgly.....ProArgArgPr 13
572 AACCATCTCGGTGCACTGACAGCTCTGAGTACTTAAGCGGAC 523
    ||||||| || ||||| |||
13  OArgProProGlyProThrGlyAlaValProThrAlaGlnSerGlnValT 30
522 GCAGCTCTCTCAGAGAGAGCCGCT.....GCCGATGCGTTCGTACG 479
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
30  hSerThrProAsnSerGlnProAlaValArgSerAlaProAlaAlaAla 46
478 CCACCTCCGCCGACGAGTGCACATATGATGAGTCCGCTCGCAGTGA 429
    |||||||
47  ProProPro.....ProProAlaGly.. 53
428 ATGGACTCTTCGATGACTGTGACACTCTCCAGCGATTGTACATGC. 380
    ||||||| ||||| |||||
54  .....GlyProProProSerGlySerLeul. 62
380 ..... 380
62  euleuArgGlnTrpLeuHisValProGlnSerAlaSerAspAspAsp 78
379 .....CTGCTCCAGCGAGCCG 363
79  AspAspAspTrpProAspSerProProGlnSerAlaProGlnAlaAr 95
362 TTTCGTCGACATGTTCTCCAGCAAGCATACGAGCGCTAAATAATGAGCG 313
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
95  gProThrAlaAlaAlaProArgProProGlyProHis..... 107
312 CTCCTATTATTAGCTTCTTGTGGTGGGGGGGTGCGACTGCAGCA 263
107 ..... 107
262 TGTCGACCGTACATACTGAAAGGCGATGGCCACTTCTTGGGATCT 213
    ||||| ||||| |||||
108 .....ArgProAlaTrpAlaArgGlyAlaGlyLe 117
212 TTGTACATCAGCTCCAGAGATTACCCGCTTGACGTTGCCCACTTCGC 163
    ||||| ||||| |||||

```

```

117  uThrProThrProProAlaPro.....SerAlaPhea 130
162 GAGGCGCTCGGTACAGCAAGATGCTGTTGTACTTCTCGAAGGT 113
    ||||||| ||| ||| |||
130  rArgAla...SerProSerAlaCysAlaSerProArgSer..... 142
112 TCAATACGTGTCTCCCGCTGCGGATGTTGCCCTGATGAGACGCTG 63
    ||||||| ||| ||| |||
143  .....ThrTrpAlaAla...CysAlaCysAspAlaArgAlaGlyArg 156
62  AAGGCTGGGTGCGCTCGGCTGACTGTGTCGCCCTTTCGGGACAGATGC 13
    ||| ||| |||||
156  yArgArgSerProProArgProProProProProProProProPro 173
12  TGCCTCGCCCA 2
173  roArgProPro 176

```

seq_name: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:PCT-US91-06532-3

seq_documentation_block:

```

; Sequence 3, Application PC/TUS9106532
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; TITLE OR INVENTION: Recombinant Herpes Simplex Viruses
; TITLE OF INVENTION: Vaccines and Methods
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: Two First National Plaza Suite 2100
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/06532
; FILING DATE: 19910910
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Grubert, Lewis S.
; REGISTRATION NUMBER: 30,060
; REFERENCE/DOCKET NUMBER: 27373/8235
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/346-5750
; TELEFAX: 312/984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US91-06532-3

```

alignment_scores:
 Quality: 98.00 Length: 237
 Ratio: 1.210 Gaps: 10
 Percent Similarity: 34.177 Percent Identity: 23.629

alignment_block:

US-09-416-579a-1/rev x PCT-US91-06532-3 ..

Align seg 1/1 to: PCT-US91-06532-3 from: 1 to: 355

```

622 CGGCATCGAGACTAGACCTTCAGACACTGGCGTCTGTGGTATC 573
    |||||||

```

```

6  ArgHisArggly.....ProArgArgpr 13
572 AACGAGTCGTCGACATCATGCTCCTGAGGTACTTAAGCGGCAC 523
    ||||| ||| ||||| |||
13  cArgProGlyProThrGlyAlaValProThrAlaGlnSerGlyValr 30
522 GCAGCTCTCTCAGACAGAGCCGCT.....GCCGATCCGCTTCGTACG 479
    ||||| ||||| ||||| |||||
30  hrSerThrProAsnSerGluProAlaValArgSerAlaProAlaAla 46
478 CCACCTCCGGGAGGTGGCAGATATATGATGAGTCCGCTCAGGTGA 429
    ||||| ||||| ||||| |||||
47  PropPro.....ProProAlagly.. 53
428 ATGAGCTCTGATGACTTGTACACTCCTCCAGCCTATGTTACATGC. 380
54 .....GlyProProSerSerLeul 62
380 ..... 380
62  euleuArgGlnTrpLeuHisValProGluSerAlaSerAspAsp 78
379 .....CCTGCTCCAGGAGCCG 363
79  AspAspAspTrpProAspSerProProGluSerAlaProGluAlaAr 95
362 TTTCGTCGATGTTCCGACGAGCATATAGCGGCTAAATGAGCG 313
    ||||| ||||| ||||| |||||
95  gProThrAlaAlaAlaProArgProProGlyProHis..... 107
312 CTCCTTTATTTTACCTCTCTGTTGGTGGCGGTGTCGACCTGACCA 263
107 ..... 107
262 TGGTCAGCGTGAATACCTGTAAGGCGATGCGCCACTTCTGGGATCT 213
    ||||| ||||| ||||| |||||
108 .....ArgProAlaTrpAlaArgGlyAlaGlyLe 117
212 TTGACATCAGCTCCAGACATTTACCCGCTGAGCGTTCGACCTTC 163
    ||||| ||||| ||||| |||||
117  uThrProThrProArgProAlaPro.....SerAlaPheA 130
162 GAGCGCTCGCTCAGCAGCAATGCTGTTGTTACTTCTCGAAGTGT 113
    ||||| ||||| ||||| |||||
130  rGArgAla...SerProSerAlaCysAlaSerProArgSer..... 142
112 TCATATACGTGCTCTCCCGCTGCCGATGTTGCCCTCGATGAGAGCG 63
    ||||| ||||| ||||| |||||
143 .....ThrTrpArgAla...CysAlaCysAspAlaArgAlaGlyArg 156
62  AAGGCTGGTGCCCTCGGCTACTTGGTCCCTTTCGGCAGCAGATGC 13
    ||||| ||||| ||||| |||||
156  yArgArgSerProArgProArgProArgProArgProArgProArg 173
12  TGCCCTCCGCA 2
    ||||| ||||| ||||| |||||
173  roArgProPro 176

```

seq_name: /cgn2_6/ptodata2/1aa/5H_COMB.pep:us-08-660-963-12

seq_documentation_block:

```

: Sequence No. 5852187
: Patent No. 5852187
: GENERAL INFORMATION:
: APPLICANT: Thornton, Michael O.
: APPLICANT: Gaylinn, Bruce D.
: APPLICANT: Horikawa, Reiko
: APPLICANT: Lyons Jr., Charles E.
: TITLE OF INVENTION: MOLECULAR CLONING OF THE OVINE PITUITARY
: GROWTH HORMONE RELEASING HORMONE RECEPTOR
: NUMBER OF SEQUENCES: 18
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: POPHAM, HAIR, SCHNOBRICH & KAUFMAN, LTD.

```

```

? STREET: Metropolitan Square Building, Suite 800, 1450
? STREET: G. Street
? CITY: Washington
? STATE: D.C.
? COUNTRY: USA
? ZIP: 20005
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? FILING DATE: 12-JUN-1996
? CLASSIFICATION: 536
? ATTORNEY/AGENT INFORMATION:
? NAME: O'Shaughnessy, Brian P.
? REGISTRATION NUMBER: 32,747
? REFERENCE/DOCKET NUMBER: 18046.036
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-824-8199
? TELEFAX: 202-824-8199
? TELEX: 248516
? INFORMATION FOR SEQ ID NO: 12:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 498 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-660-963-12

```

alignment_scores: Quality: 96.50 Length: 231
 Ratio: 0.995 Gaps: 10
 Percent Similarity: 41.991 Percent Identity: 24.242

alignment_block:
 US-09-416-579a-1/rev x US-08-660-963-12 ..

Align seg 1/1 to: US-08-660-963-12 from: 1 to: 498

```

604 CCTTGACAGACGTCGCGTCGCTGATACACAGTCTGTGACAC 555
    ||| ..... ||||| ||||| ||||| |||||
191 ProSerSerTrpSerLeuSerGlyGlySerThrAlaProGlyThrTh 207
554 TCATGACAGCTCTGAAGTACTTAAGCGGACGCGAGCTCTCTCAGAACG 505
    ||| ..... ||||| ||||| ||||| |||||
207 r.....SerThrProSerCys 213
504 AGCCCGCTGCCGA.....TGCGTGTACGCGACCT... 473
    ||||| ..... ||||| ||||| ||||| |||||
213 erProProLeuSerSerArgArgGlnLeuCysSerArgThrProProSer 229
472 .....CGGCGAGGTGCCGAGA 456
230 PheThrGlyArgThrTrpThrAlaAlaSerProLeuSerCysAlaAr 246
455 TATATGATGAGTCCGCT.....CGAGGTG 430
    ||||| ..... ||||| ||||| ||||| |||||
246 gLeu.....LeuProProLeuIleSerArgProProThrSerAlaGlyC 261
429 AATGAGCTCTCGATGACTTGTACACCTCCGAGGATTTGT..... 386
    ||| ..... ||||| ||||| ||||| |||||
261 yStrpGlnIleLeuSerThrProAlaSerProPro...HisCysProAla 276
385 .....ACATG 381
277 GlnGlyGlySerSerGlyGlyTrpPheSerLeuProGlyGlyPheLeu 293
380 CCCTGCTCCAGGAGCGCGTTCGTCGATGTTCTCCACGAAATAGCG 331
    ||||| ..... ||||| ||||| ||||| |||||

```

us-09-416-579a-1.rai

US-08-209-521-11

```

alignment_scores:
  Quality: 95.50
  Ratio: 1.027
Percent Similarity: 44.076
Percent Identity: 21.801

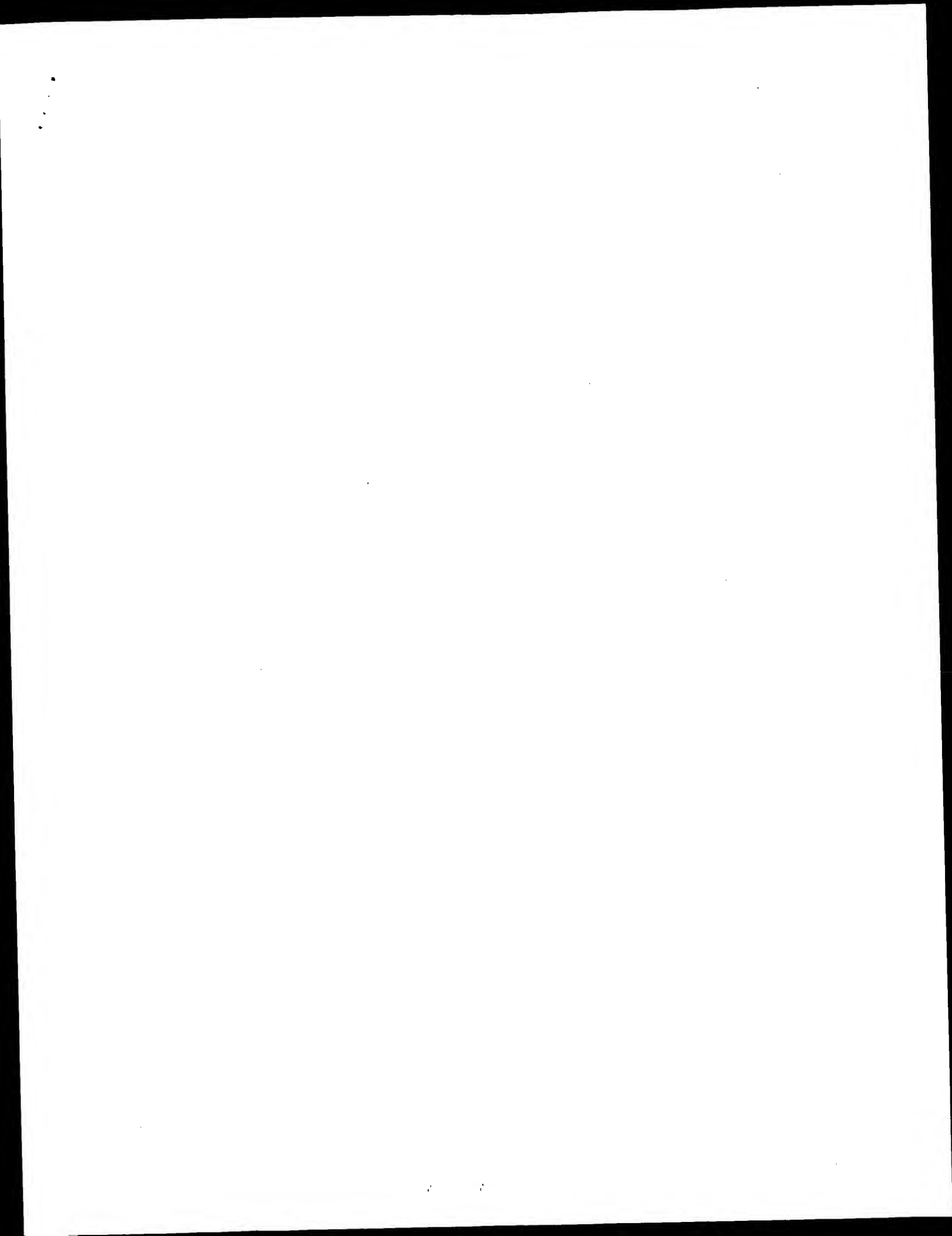
```

alignment_block:

US-08-209-521-11

Align seg 1/1 to: US-08-209-521-11 from: 1 to: 341

603 TGGCGCTCCGACGCGGTGACTCCGGGC...CAATGTTTCCAGGTTG 627
132 Cyscysalaaglythralthralthrthrthrcysalthracyscysthralthr 148
626 AGATCGGCATCGAGACCTAGGACCTTCACAGCTGGCGGTCGTCTGGTG 577
148 T.....glyelcysthrthrthrthrcysglyalaglcy 159
576 TATCA.....ACCAGTCTGGTGCATCATGAGCTCCTCGAA 539
159 lythrthrglyalaglyglythralthralaglcythralthraalaglalya 175
538 GGTACTTAACGGCAGCGAGCTCTCCTCGAAGACCCGCTGGCGGA.. 491
176 ThrthrcysalalaglalyalalalathrthrglythrlyThralthralal 192
490TGCGTGGTGGCCACCTCCGGCCA 466
192 aalathralthrthrcyscysthrthrcyscysthrthrglyalathrthrgly 209
465 GGTGCCCATATATGATGATGAGTCCGCGCTCGAGGTGAATGAGACTCTTGC 416
209 lacysla.....ThrthrglyThrthyrthrthrglythr 219
415 TGAACCTTGACCACTCTCCAGCGATGTGATCATGCCCTGGTCCAGCGAG 366
220 cysalathrthrthrglythrthralaglythralthralthrglythralthr 236
365 CGGTTTCGCGCATGTCTCCAGCAAGCAATAGCGAGCGCTAAATAATGA 316
236 T..... 236
315 GCGCTCYTATTTTACTTCTGTGGTGGGGCGGTGCGACATCGA 266
237ThrrhrcysthrCysalalacysalathrAla 247
265 GCATGTGACGTGATCAATACCTGAAAGGCGATGCGCCACTTGTGGGA 216
248 Glyalathralthraalalathralthralaglalythrthrthrglythr 264
215 TCTTGTGATCAGCTCCAGCAGAGATTACCCCGTGGACGTTGGCGCACTT 166
264 rAlacyscysthrthrrthr.....Alacysthrthrg 276
165 CTCGACGGCGCTCGGTCAAGCGCAAGATGTCGTTCTTACTTCCGCAAGT 116
276 lythrthralalalalathrthrglythralthrthrglycysalalalathr 292
115 GGTCAATATACGGGTCTTCCGCGCTCCGCAATG 83
115 GlythralthrthrglycysalalalalathrCys 303




```

|||||
17 ythnglnpophethValleuIIeGLuGIYasnIIeclYserGIYst 34
101 CCACGATTTGAACACCTTCGAGAGTACAGAGACGACATTTGCTGCTG 150
34 hrtHrYrleuasnHIsphneGIuYrYLYsasnHIIeclYserleu 50
151 ACCGAGCGCTCGAGAAGTGGCGCAACGTCAACGGGGTAAATCTGCTGCA 200
51 ThnglnpovaIIuIuYstrpAtgaenValasnGIYValasnleuGI 67
201 GCTGATGTACAAAGATCCCAAGAAAGTGGCCATGCCCTTCAGATTATG 250
67 uIeMeTtYrYsAspYrYsTrpAlaMetProPheGlnserYrY 84
251 TCACGGTACCATGCTGACGTCCACACCGCCCAACACCAAGAAAGCTA 300
84 aIrrHleuThMeIeugInserHIsHrAlaProThrasnYsYsleu 100
301 AAAATAAAGAGCGCTCCATTTTACGCGCTGCTATGCTGCTGAGAA 350
101 LysIleMeTGIuArYserIlePheSerAlaArYrYcYsPheValGIuAs 117
351 CATGGAGCAAAACGGCTGCTGAGCAGCGCATGTACATACGCTGAGC 400
117 nMeTArYgaSnGIYserIeugInGIYMeTtYrYasnThleuGIu 134
401 AGTGTACAACTTCATCGAAGATCCATTCACCTGACGGGAGCGCTCAG 450
134 lUTrpYrYsPheIIeGIuInserIleHIsValGIuAlaAsPheIIe 150
451 AATATATCTGCGACCTCGCGGAGGTGGCTACGACGATCGCGGAGCG 500
151 lIeYrYleuArYThSerProGIuValAlaArYrYlArYlArYGIuAr 167
501 GGCTGCTGTGAGAGAGCTGCGCTCGCTTAAGTACCTTACGAGCTGC 550
167 gAlaArYserGIuInserCYsValProleuYsTrYrYleuInleuH 184
551 ATGAGTTGCACAGACGACTGTTGATACACAGACAGACCGCAGTCTGC 600
184 lsgIuIeunHIsGIuAspTrPheIIeHIsGIuArYArYProGIuInserCYs 200
601 AAGCTCTAGTCTCGATGCCATCTGAAACCTGGAACATTTGGACCGGA 650
201 LysValleuValleuAspAlaAsPheuAsnleuGIuAsnIIeGIYThnGI 217
651 GTACACAGCGCTCGAGAGCAGCATATTTCAGCGCATCTCAAGTAACAA 700
217 uYrYleuInArYserIlePheAspAlaIleSerSerAsnGIu 234
701 AGCCCTCGCGGCTTCTGTGTGCTGCCCAAGCAAGCGCGGAGGGCGCAGA 750
234 lnpProserProValleuValserProserYsArYGIuArYAlaAlaArY 250

seq_name: /SID8/gcgdata/geneseq/AA2000.DAT:AA20726
seq_documentation_block:
ID   AAG20726 standard; protein: 420 AA.
XX
XX   AAG20726;
AC
XX
XX   17-OCT-2000 (first entry)
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 23205.
XX
XX   Protein identification: signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
XX   termination sequence.
OS   Arabidopsis thaliana.
XX
XX   EPI033405-A2.
*PN

```

```

XX   06-SEP-2000.
XX
XX   25-FEB-2000; 2000EP-0301439.
XX
XX   25-FEB-1999; 99US-0121825.
XX   05-MAR-1999; 99US-0123180.
XX   09-MAR-1999; 99US-0123548.
XX   23-MAR-1999; 99US-0125788.
XX   25-MAR-1999; 99US-0126264.
XX   29-MAR-1999; 99US-0126785.
XX   01-APR-1999; 99US-0127462.
XX   06-APR-1999; 99US-0128334.
XX   08-APR-1999; 99US-0128714.
XX   16-APR-1999; 99US-0129845.
XX   19-APR-1999; 99US-0130077.
XX   21-APR-1999; 99US-0130510.
XX   23-APR-1999; 99US-0130891.
XX   28-APR-1999; 99US-0131449.
XX   30-APR-1999; 99US-0132048.
XX   04-MAY-1999; 99US-0132407.
XX   05-MAY-1999; 99US-0132484.
XX   06-MAY-1999; 99US-0132485.
XX   06-MAY-1999; 99US-0132486.
XX   07-MAY-1999; 99US-0132487.
XX   11-MAY-1999; 99US-0132863.
XX   14-MAY-1999; 99US-0134256.
XX   14-MAY-1999; 99US-0134218.
XX   14-MAY-1999; 99US-0134219.
XX   14-MAY-1999; 99US-0134221.
XX   18-MAY-1999; 99US-0134370.
XX   19-MAY-1999; 99US-0134941.
XX   20-MAY-1999; 99US-0135124.
XX   21-MAY-1999; 99US-0135353.
XX   24-MAY-1999; 99US-0135629.
XX   25-MAY-1999; 99US-0136021.
XX   27-MAY-1999; 99US-0136392.
XX   28-MAY-1999; 99US-0136782.
XX   01-JUN-1999; 99US-0137222.
XX   03-JUN-1999; 99US-0137502.
XX   04-JUN-1999; 99US-0137502.
XX   07-JUN-1999; 99US-0137724.
XX   08-JUN-1999; 99US-0138094.
XX   10-JUN-1999; 99US-0138540.
XX   10-JUN-1999; 99US-0138847.
XX   14-JUN-1999; 99US-0139119.
XX   16-JUN-1999; 99US-0139452.
XX   17-JUN-1999; 99US-0139453.
XX   17-JUN-1999; 99US-0139492.
XX   18-JUN-1999; 99US-0139454.
XX   18-JUN-1999; 99US-0139455.
XX   18-JUN-1999; 99US-0139456.
XX   18-JUN-1999; 99US-0139457.
XX   18-JUN-1999; 99US-0139458.
XX   18-JUN-1999; 99US-0139459.
XX   18-JUN-1999; 99US-0139460.
XX   18-JUN-1999; 99US-0139461.
XX   18-JUN-1999; 99US-0139462.
XX   18-JUN-1999; 99US-0139463.
XX   18-JUN-1999; 99US-0139750.
XX   18-JUN-1999; 99US-0139753.
XX   18-JUN-1999; 99US-0139817.
XX   21-JUN-1999; 99US-0139889.
XX   22-JUN-1999; 99US-0139899.
XX   23-JUN-1999; 99US-0140353.
XX   24-JUN-1999; 99US-0140354.
XX   26-JUN-1999; 99US-0140695.
XX   28-JUN-1999; 99US-0140823.
XX   29-JUN-1999; 99US-0140981.
XX   30-JUN-1999; 99US-0141287.
XX   01-JUL-1999; 99US-0141842.
XX   01-JUL-1999; 99US-0142154.

```

```

PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145108.
PR 23-JUL-1999; 99US-0145109.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147360.
PR 06-AUG-1999; 99US-0147760.
PR 06-AUG-1999; 99US-0147703.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.

```

```

PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161400.
PR 25-OCT-1999; 99US-0161401.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

alignment_scores: quality: 315.50 length: 279
 ratio: 1.960 gaps: 8
 Percent Similarity: 57.706 Percent Identity: 29.032

alignment_block:
 US-09-416-579a-1 x AAG20726 ..

Align seg 1/1 to: AAG20726 from: 1 to: 420

```

49 GGCACCCAGCC.....TTCACGCTCTCATCGA 77
   ||| :|||
95 GYPRGGLIPIROHASNLEUASVALLYSARGLEUTHRPHCYSAIGI 111
78 GGCACATCGGCACGCGGAGACCATTTG.....AACC 115
   ||||| :|||
111 ucllysnlleserValglYlYsserThrpheLengInAArgIleAlaAsng 128
116 ACTTCGAGAGTACAGACAGACATTTGCTGCTGACCGAGCCCGTGA 165
   :|| :|||
128 LuThValgluLeugInspLeuValgluLeValProgluProValasp 144
166 AAGTGGCCAGCATC.....AACGGGTAAATCTGCTGAGCTGATGA 209
   ||||| :|||
145 LysTrpGlnspsValglYProAspHlsPheAsnIleLeuAspAlaPheTy 161
210 CAAGATCCCAAGAGTGGCCATTCCTTTACAGATTATGTCAGCTGA 259
   |:::| :|||
161 rSecluproGlnArgTyAlaTyThrpheGlnsnyTyAlaPheValT 178
260 CC.....ATGCTGACGTCACACCGCCCAACCAACAGAGCTAA 303
   ||| :|||
178 hrArgLeuMetGlnGluTyGlnSerAlaSerGlyValLysProLeuArg 194

```

[illegible]


```

PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147312.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148665.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

```

```

PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

alignment_scores:

```

Quality: 315.50 Length: 279
Ratio: 1.960 Gaps: 8
Percent Similarity: 57.706 Percent Identity: 29.032

```

alignment_block:

```

US-09-416-579a-1 x AMG20725 ..

```

```

Align seg 1/1 to: AMG20725 from: 1 to: 523

```

```

49 GGCACCCAGCC.....TTACCCGTCATCGA 77
198 GTPGluProAsnLeuAsnValLysLysArgLeuThrPheCysValIgl 214
78 GGCACATCGCGCAGCGGAGACCCGATTTG.....AACC 115
214 uGlyAsnLleSerValIglLysSerThrPheLeuGlnArgIleAlaAsnG 231
116 ACTTCGAGAGTACAGACGACATTTGCCCTGACCGACCCGCTCGAC 165
231 lUThrValGluLeuGlnAsnLeuValIglIleValIProGluProValAsp 247
166 AAGTGGCCGACGTC.....ACGGGTAATCTGCTGAGCTGATGA 209
248 LysTrpGlnAspValIglLysProAspHisPheAsnIleLeuAspIlePhe 264
210 CAAGATCCCAAGAGTGGCCATGCCCTTTCAGAGTATGTCAGCTGA 259
264 rSerGluProGlnArgLysIleValIThrPheGlnAsnTyValIlePheVal 281
260 CC.....ATGTCGACATCCGACACCCGCCCAACCAAGAGCTGA 303
281 hrArgLeuMetGlnGluLysGlnSerIleValIleValIleValIleVal 297
304 ATAAAGAGCGCTCATTTTACGCGCTGCTGCTGCTGCTGCTGCTGCTG 353
298 LeuMetGlnArgSerValIlePheSerAspArgMetValIleValIleVal 314
354 GCGACGAAGCGCTCGCTGAGCGAGGAGATGACATACGCTGAGAGT 403
314 lHisGluAlaLysTrpMetAsnGluMetGluIleSerIleLysIlePhe 331
404 GTTACAGTTCATCGAAGAGTGCATT...CACCTGACGGGAGCTCATC 450

```



```

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148177.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

```

```

PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160774.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

alignment_scores:

Quality:	315.50	Length:	279
Ratio:	1.960	Gaps:	8
Percent Similarity:	57.706	Percent Identity:	29.032

alignment_block:

US-09-416-579a-1 x AAG20724

Align seg 1/1 to: AAG20724 from: 1 to: 580

```

49 GGGACCCAGGCC.....TTACCGTCCTCATGCA 77
255 GTPProgluProbsnleuansvallylsyargLeuthrPheCysVal 271
78 GGGCAACATCGGAGCGGAGAGACCGATTTG.....AACC 115
271 uclYasnIleSerValIGlyLysSerThrPheLeuGlnArgIlealaang 288
116 ACTTCGAGAACTACAGAAAGACATTCGCTGACCGACCCGTCGAG 165
288 IuthrValIGluLeuGlnAspLeuValIGluIleValProgluProvalAsp 304
166 AAGTGCGCGACGTC.....AACGGGTAATCGCTGGAGCGATGTA 209
305 LysTrpGlnAspValIGlyProAspHisPheAsnIleuAspAlaPheTy 321
210 CAAGATCCCAAGAGAGTGGCCATGCCCTTACAGATTATGTCACGCTGA 259
321 fserIuProGlnArgYrAlaTyThrPheGlnAsnTy-ValIleVal 338
260 CC.....ATGCTGAGTCGCGACACCGCCCAACCAAGAGCTAA 303
338 hrArgLeuMetGlnGlnIuysGlnSerIaSerGlyValIysProLeuArg 354
304 ATAAAGCAGCGCTCCATTTTACGCTGCTATTCCTTCGAGAAACAT 353
353 LeuMetGlnArgSerValIlePheSerAspArgMetValPheValArgAla 371
354 GCGAGAAACGCGCTCGCTGGAGCGGACATGTACATACCGTCGAGAGT 403
371 IHisGlnAlaIleStrpMetAsnGlnMetGluIleSerIleTyAspSer 388
404 GGTACAGATTATCATGAGAGTCCAT...CACCTGACGCGGACCTCATC 450
388 rPheAspProValIleSerSerLeuProGlyLeuValIleProAspGly 404
451 ATATATCTGGCACTCCGCGGAGGTCGGCTACGAGACGATCCCGACGG 500
405 IleTyTrpLeuArgIleAspProAspTrpIleCysHisLysArgMetLeuArg 421
501 GGTCTGTTCTGAGAGAGTGCCTGCGCTTAAGTACCTAGACCTGCGCTGC 550
421 gLysArgAlaGlnGluAspGlyValSerLeuTyIleuGlnAspLeuH 438

```

```

551 ATGAGTTCACACGAGCTGTTG..... 573
||||| |||:|||||
458 ISGLIVSHISGLISERTIPRELEUDEPROHEGLISERGLYASNHSGLY 454
574 ..ATACACGAGAGCGACG..... 591
|||||
455 VALLEUSERVALSERARGPROSERLEUHHISMETASPANSERLEUHSIPR 471
591 ..... 591
471 CASHLELYSASPARVALPHETRYLEUGLGLYASNHSMETHSERS 488
592 ....CAGTCGCAAGTCCAGTCCGATGCCGATCGACCTGAA 636
||||| :|||||:|||||:|||||
488 ERLLEGLINLYSVALPROALAUEUVALDEUSPCYSGLUPROASNILEASP 504
637 .....ACATTGGCACCGAGTACCGACGCTCGAGAGACGAG 671
||||| :||||| :|||||
505 PHESERARGAPRILEGLUALALYSTHRCINTYRALAARGINVALAIGL 521
672 CATATTGACGCCATCTCAAGTAAACCAACAGCCCTCG 708
:|||||: :|||||:|||||:|||||
521 UPHEPEGIUPHEVALLYSLYSGLINLTHRSER 533
seq_name: /SIDS8/gcdata/geneseq/AA1998.DAT:AAW70493

```

seq_documentation_block:

ID AAW70493 standard: Protein: 261 AA.

AAW70493:

29-DEC-1998 (first entry)

Human disease related nucleotide kinase-1 (DRNK-1) protein sequence.

Human disease related nucleotide kinase-1; DRNK-1; deoxyguanosine kinase;

p21ras; cell proliferation; oncogenesis; cancer; gene therapy;

Immune disorder; neurological dysfunction.

Homo sapiens.

Key Location/Qualifiers

peptide 1..25 /note= "Mitochondrial localisation signal"

Modified-site 8 /note= "N-myristoylated"

Domain 28..34 /note= "ATP/GTP binding site (P-loop)"

Modified-site 141 /note= "N-glycosylated"

US5817482-A.

06-OCT-1998.

20-JUN-1997: 97US-0879561.

20-JUN-1997: 97US-0879561.

(INCY-) INCYTE PHARM INC.

Bandman O, Corley NC, Guegler KJ, Hawkins PR, Hillman JL;

WPI; 1998-556388/47.

N-PSDB; AAV33481.

Nucleic acids encoding deoxyguanosine kinase - useful for

recombinant production of the enzyme for treating diseases caused by

lack of the enzyme e.g. cancers caused through loss of enzyme

function

```

XX The present sequence represents a human disease related nucleotide
CC kinase-1 (DRNK-1) protein sequence. The DRNK-1 encoding DNA sequence
CC was first identified in Incyte Clone 56821 from the fibroblast cDNA
CC library (FIBRNOT 01). The DRNK-1 DNA sequence is useful for the
CC production of the corresponding recombinant enzyme. The invention
CC provides DRNK enzymes, which are deoxyguanosine kinases, which catalyse
CC the transfer of a terminal phosphate from adenosine triphosphate (ATP)
CC or guanine triphosphate (GTP) to guanosine or guanine diphosphate.
CC As GTP levels are known to control the activity of certain oncogenic
CC proteins e.g. p21ras, a protein involved in cell proliferation and
CC oncogenesis, suppression of the enzyme activity causes high ratios of
CC GTP:GDP, promoting oncogenesis. Therefore, diseases (e.g. cancers,
CC immune disorders and neurological dysfunction) related to this lack of
CC activity may be prevented or treated with the recombinant enzyme, or by
CC gene therapy based strategies. Anti-sense constructs of the DRNK
CC encoding nucleic acids may also be used for inhibition of over-expression
CC of the enzyme.
SQ Sequence 261 AA:

```

alignment_scores: Quality: 287.50 Length: 246
Ratio: 2.054 Gaps: 7
Percent Similarity: 56.911 Percent Identity: 29.268

alignment_block:
US-09-416-579a-1 x AAW70493

Align seq 1/1 to: AAW70493 from: 1 to: 261

```

49 GGCACCCAGCCCTTCCACGTCCTCATCGAGGCAACATCGCAGCGGGA 98
||| ||| :|:|||||:|||||
18 G|A|A|G|L|Y|P|R|A|A|G|L|Y|S|E|R|I|L|E|G|L|Y|A|S|N|I|L|E|A|V|A|G|L|Y| 34
99 GACCACGATATTGACCACTTCGAGAGTACAAAGAACGATTTCCCTGC 148
:|||||: :|:|||||:
34 S|S|E|R|T|H|P|H|E|V|A|L|Y|S|L|E|U|E|H|T|H|Y|R|T|Y|R|P|R|O|G|I|U|T|R|P|H|S|V|A|L|A 51
149 TGACCGAGCCCGTCGAGAAAGTGCGCAGACGTCACAGCGGGA..... 189
|||||:|||||:
51 L|A|T|H|G|L|I|P|R|O|V|A|L|A|L|A|T|H|T|R|P|G|L|A|S|N|I|L|E|G|L|A|L|A|G|L|Y|T|H|G|L|N 67
|||||:|||||:
190 .....AATCTGCTGAGCTGATGATGACAA 212
|||||:|||||:
68 L|Y|S|A|L|A|C|Y|T|H|A|L|A|I|N|S|E|R|L|E|U|G|L|Y|A|S|N|L|E|U|D|E|A|S|P|H|E|T|Y|R|A|T 84
:|||||: :|:|||||:
213 A|G|A|T|C|C|A|A|G|A|G|G|G|C|A|T|G|C|C|T|T|C|A|G|A|C|T|T|A|T|G|T|C|A|C|G|T|G|A|C|A 262
:|||||: :|:|||||:
84 G|L|U|P|R|O|A|L|A|A|T|G|T|P|S|E|R|Y|R|T|H|P|R|H|E|G|I|N|T|H|P|H|E|S|E|R|P|H|E|U|S|E|R|A 101
263 T|G|C|T|G|C|A|G|T|C|G|C|A|C|G|C|C|C|A|C|A|G|A|A|G|T|A|..... 300
|||||: :|:|||||:
101 R|G|L|E|U|L|Y|S|V|A|L|G|L|N|E|U|G|L|U|P|R|O|P|R|H|E|P|R|O|G|L|U|Y|S|L|E|U|G|L|A|L|A|A|T|G 117
|||||: :|:|||||:
301 .....AAATTAARGAGCGCTCCATTTTATGCTGCTATTCGTT 341
:|||||: :|:|||||:
118 L|Y|S|P|R|O|V|A|L|G|I|N|L|E|P|R|H|E|G|L|U|A|R|G|S|E|R|V|A|L|Y|S|E|R|S|P|R|A|T|G|L|E|P|R|H 134
|||||: :|:|||||:
134 E|A|L|A|Y|S|A|S|N|L|E|U|R|H|E|G|L|U|A|S|N|G|L|Y|S|E|R|L|E|U|S|E|R|A|S|P|..... 146
392 C|G|T|G|G|A|G|A|G|T|G|T|A|C|A|A|G|T|T|C|A|T|G|C|A|G|A|G|T|C|A|C|T|G|C|A|G|G|C 441
|||||: :|:|||||:
147 .....I|E|S|E|R|S|E|R|G|L|Y|I|L|E|S|E|R|I|L|E|A|R|G|T| 155
442 G|A|C|C|T|G|A|T|G|A|T|A|.....T|A|T|C|G|G|C 461
|||||: :|:|||||:
156 G|L|Y|I|L|E|U|R|P|H|E|S|E|R|C|Y|S|G|L|Y|S|E|R|L|E|U|R|P|R|O|A|L|G|L|Y|S|E|R|H|I|S|T|Y|R|M|E|T|A|I 172

```

10


```
seq_name: /STDS8/gcgdata/geneseq/geneseq/AA1999.DAT:AAV31745
seq_documentation_block:
ID      AAV31745 standard: Protein; 430 AA.
XX
XX
AC      AAV31745;
DT      22-NOV-1999 (first entry)
XX
DE      Mycobacterium tuberculosis specific DNA-encoded polypeptide
XX
XX      Tuberculosis; infection; diagnosis; DNA probe.
XX
OS      Mycobacterium tuberculosis.
XX
XX
Key      Location/Qualifiers
FH      Misc-difference 4
FT      /note= "encoded by TGA"
FT      Misc-difference 6
FT      /note= "encoded by TGA"
FT      Misc-difference 20
FT      /note= "encoded by TGA"
FT      Misc-difference 29
FT      /note= "encoded by TGA"
FT      Misc-difference 54
FT      /note= "encoded by TGA"
FT      Misc-difference 64
FT      /note= "encoded by TGA"
FT      Misc-difference 69
FT      /note= "encoded by TGA"
FT      Misc-difference 89
FT      /note= "encoded by TGA"
FT      Misc-difference 99
FT      /note= "encoded by TGA"
FT      Misc-difference 114
FT      /note= "encoded by TGA"
FT      Misc-difference 119
FT      /note= "encoded by TGA"
FT      Misc-difference 129
FT      /note= "encoded by TGA"
FT      Misc-difference 159
FT      /note= "encoded by TGA"
FT      Misc-difference 169
FT      /note= "encoded by TGA"
FT      Misc-difference 182
FT      /note= "encoded by TGA"
FT      Misc-difference 185
FT      /note= "encoded by TAG"
FT      Misc-difference 219
FT      /note= "encoded by TGA"
FT      Misc-difference 259
FT      /note= "encoded by TGA"
FT      Misc-difference 269
FT      /note= "encoded by TGA"
FT      Misc-difference 291
FT      /note= "encoded by TGA"
FT      Misc-difference 323
FT      /note= "encoded by TGA"
FT      Misc-difference 339
FT      /note= "encoded by TGA"
FT      Misc-difference 349
FT      /note= "encoded by TGA"
FT      Misc-difference 356
FT      /note= "encoded by TGA"
FT      Misc-difference 366
FT      /note= "encoded by TGA"
FT      Misc-difference 410
FT      /note= "encoded by TGA"
FT      Misc-difference 421
FT      /note= "encoded by TGA"
FT      /note= "encoded by TGA"
XX      EP945462-A1.
XX
```


CC The present invention relates to toluene degrading enzyme genes and
CC proteins *tutH* (see AAF33629 and AAB59831), *tutI* (AAF22630 and AAB59832),
CC *tutF* (AAF22631 and AAB59833) and *tutG* (AAF22632 and AAB59834). The
CC toluene degrading enzymes are homologues of pyruvate formate lyase. The
CC toluene degrading enzymes are useful for biological treatment of organic
CC compounds and in particular for the degradation of toluene and its
CC analogs contained in liquid or solid waste source. The present sequence
CC is a protein sequence for toluene degrading enzyme, *TutD*.
XX
XX
XX Sequence 1017 AA:

alignment_block:
HS-09-416-E707

US-09-416-579A-1 x AAB59813

.....: seg 1/1 CO: AAB59813 from: 1 to: 1017

```

28  |||||||:.....:|||||
29  TrrpAgtAgtArgProSerLeuProValThrThrAlaSerCysArgProth 91
30  .....
31  91  LysProSerTrpLysThrGlyCysTrpParGalaSerSerProLys 108
32  CACCACGCCCTTACCCTCC.....
33  :|||:|||||:
34  108  erLleSerProLysProArgProThrcysArgProGlyThrAla 124
35  71  .....
36  125  ArgArgValSerThrThrSerProArgSerThrThrGlyArgArgTrpse 141
37  73  ATGAGGGCCACATCGCGACGGCGAAGACCAACGATTGTGAACCACTTCGA 122
38  |||||  |||  :|||:|||||:
39  141  rSerProAlaArgSerAlaGlyArgAlaGlyArgAlaGlyCysAla 158
40  123  GAAGTCAGAAAGACGACATTTGGCTGCTGCTGACCGAGCCCGTCGAGAAGTGG 172
41  |||||:||||  |||  |||:|||||:
42  158  rGserSerTrpGlyStr.....SerArgProLleArgSerAla 170
43  173  GCACAGCTCAACGGGTAAATCATCGTGGAGCTGARGTCAAAAGTCCCAAG 222
44  |||  |||:||||  |||
45  171  ArgProSer.....CysSerLysSerProth 179
46  223  AGGT.....GGCCATGCCCTTTACAGATT..... 247
47  |||  |||  |||  |||:|||||:
48  179  rSerValSerAlaPheProProSerProAlaArgAlaSerAlaGlyThrArgC 196
49  247  .....
50  196  yArgArgAraSerLeuProSerSerValThrArgSerSerAlaThrArg 212
51  248  .....ATGTCACGCTACCATGCTGGAGTGCACA.....CC 279
52  213  AlaAlaThrProArgGlyStrProCysGlyArgThrThrArgPr 229
53  280  GCCCAACCAACAAGAAGCTAAATAATARGAGCGCCCTCATTTTAGGCC 329
54  |||||:|||||:|||||:

```

... ||:::||
::||:::||
... ||:::||

ATGAGACATGGACGAACGGCTCGTGGAGCAGC 379

380 GCATGTACATACTCCGCGAGGCACGGGACCACTTTC

251ValArgpHeProSerMetAlaProLaseArqAlaProTh 264

.....G 461

462 CACCTCGCCGGAGGTGGCTACCAACCA

281 IAProThrProArgAlaArgThrProAlaThrThrProThrProSerSer 297

.....TCCGGCAGCGGGCT 504
|||:||||:|||||

505 CGTTCGAGGAGAGCTGGGTGCCATTAACTACCTTTAACATCCTTCC

314 eprOArggTThrAlaArgArgCysAlaGlyPheSerSerAlaSerA 331

[illegible]

596 CGTGCAGGTCCTAGTCTCGATGCCGATCTGAACCTCCAAATCAATTTGCGGCGC

240 *Algalgasnit*..... 351

```

52 Proser[ar]acomb-1) ::||| ||| ::::|
    ||||:::||||| ::::|
```

96 CCAACAGCCCTGCCGGTTCGTGTGTCCGCCACGAAGC 733

...cyscysalacysargprolaser 379

documentation block:

AB5002C.

4-APR-2001 (first entry)

protein #3 encoded by Tnfr/E gene.

...aromatic
...; note; wood.

Shewanella metal-reducers.

200072650-A2.

- DEC-2000.

TIN-1000
CCTV

YOH -) UNIV OHIO.

scilgano PW;

DR WPI; 2001-041080/05

DR N-PSDB: AAF23627.
 XX Composition comprising toluene degrading enzyme useful for biological
 PT treatment of organic compounds, especially for degrading toluene or its
 PT analogs
 PS Disclosures: Fig 12; 122pp; English.
 XX
 XX The present invention relates to toluene degrading enzyme genes and
 CC proteins tuH (see AAF23629 and AAB59831), tuL (AAF23630 and AAB59832),
 CC tuF (AAF23631 and AAB59833) and tuC (AAF23632 and AAB59834). The
 CC toluene degrading enzymes are homologues of pyruvate formate lyase. The
 CC toluene degrading enzymes are useful for biological treatment of organic
 CC compounds and in particular for the degradation of toluene and its
 CC analogs contained in liquid or solid waste source. The present sequence
 CC is a protein sequence encoded by toluene degrading enzyme gene, tuD/E.
 XX
 XX Sequence 1615 AA:

alignment_scores:
 Quality: 137.50 Length: 346
 Ratio: 1.058 Gaps: 15
 Percent Similarity: 37.572 Percent Identity: 23.410

alignment_block:
 US-09-416-579a-1 x AAB59826 ..

Align seg 1/1 to: AAB59826 from: 1 to: 1615

```

2 TGGCGGAGGAGCAGCATCTGCGCGAA..... 28
673 trpArgArgArgProSerLeuProValThrThrAlaSerCysArgProth 689
29 ..AGGGACCACTAGCGCGAGG 50
689 rlyProSerTrpLysThrGlyCysTrpArgAlaSerSerProLys 706
51 CACCCAGCCCTCAACGTC..... 70
706 erLleSerProLysProArgProThrCysArgProSerProGlyThrAla 722
71 .....TC 72
723 ArgArgValSerThrThrSerProArgSerThrThrGlyArgArgTrpSe 739
73 ATCGAGGCGACATCGGAGCGGAGAGCACCGTATTGACCACTTGA 122
739 rSerProAlaArgArgSerAlaGlyArgAlaGlyArgAlaGlyCysAla 756
123 GAAGTACAGAGACGACATTGCTGCTGACCGAGCCGTCGAGAGAGTGGC 172
756 rGSerSerArgLysThr.....SerArgProLleArgSerAla 768
173 GCAAGCTCAACGGGGTAATCTGCTGAGCTGATGTACAAAGATCCCAAG 222
769 ArgProSer.....CysSerLysSerProth 777
223 AAGT.....GGCCATGCCCTTTCAGAGT..... 247
777 rSerValSerAlaPheProSerProAlaArgAlaSerArgThrArgC 794
247 ..... 247
794 ysaArgArgAsnSerLeuProSerSerValThrArgSerSerAlaThrArg 810
248 .....ATGTCAGCTGACCATGCTGACCTGCACAA.....CC 279
811 AlalaThrProArgArgLysThrProCysGlyArgThrThrArgPr 827
280 GCCCAACACAGACAGAGCAAAAAAFAARGAGCGCTCATTTTACGCGC 329
827 opfSerSerThrArgAsn.....S 834
```

```

330 TCAGTATGCTCTCGAGACATCGGAGAAACGGCTGCTGAGCAGG 379
834 erSerArgAlaThrTrpMetArgTrpAsn...SerSerArgTrpAsn... 848
380 GCAGTACATACGCTGAGAGGTGACAGTACAGTTCAGAAAGTCAATT 429
849 .....ValArgPheProSerMetAlaProAlaSerArgAlaProth 862
430 CACCTGACGGCGACCTCATCATATCTGC.....G 461
862 rAlaLysSerSerArgGlyArgThrLleCysSerSerSerProSerAla 879
462 CACCTGCGCCGAGGTGCGGTACGAAACGA..... 490
879 lAProThrProArgAlaArgThrProAlaThrThrProThrProSerSer 895
491 .....TCGGGACGGGCT 504
896 ArgLInProSerGlySerAlaArgProSerProProSerSerSerAla 912
505 CGTTCGAGAGAGCTGCGTGCCTTAAGTACCTCAGAGCTGCATG. 553
912 eProArgArgThrAlaArgArgCysAlaGlyPheSerSerAlaSer 929
554 .....AGTGCACACAGACTGCTGTATACACCGAGAGAGCAGCAGT 595
929 lArThrAspSerAlaLleArgArgSerSerThrThrArgSerAlaArgSer 945
596 CGTCAGAGTCTAGTCTCTGATGCGCATCTGACCTGAAACATTCGCG 645
946 ArgArgAsnThr..... 949
646 ACCGAGTACAGCGCTGAGAGACACATATTCAGCGCATCTCAAGTAA 695
950 .ProSerSerAlaSerThrAlaThrAlaProProThrArgLysProThr 966
696 COACAGCCCTGCGCGTGTGTCGTCGCCACGACAGC 733
966 hArgLysSer...ThrCysAlaCysArgProAlaSer 977

seq_name: /sids8/gcgdata/geneseq/geneseq/AA2001.DAT:AAB59817
seq_documentation_block:
ID AAB59817 standard; protein; 999 AA.
AC AAB59817;
DT 04-APR-2001 (first entry)
DE TuD protein #8.
KW toluene degradation; enzyme; waste degradation; TuD.
OS Thauera aromatica.
OS Xanthomonas maltophilia.
OS Geobacter metallireducens.
OS Azococcus toluyticus.
PN WO200072650-A2.
PD 07-DEC-2000.
PF 24-MAY-2000; 2000WO-US14298.
PR 01-JUN-1999; 99US-0323872.
PA (UYOH-) UNIV OHIO.
PI Coschigano PW;
DR N-PSDB; AAF23625, AAF23627.
```

```

alignment_scores:
  Quality: 132.50      Length: 294
  Ratio: 1.123        Gaps: 17
  Percent Similarity: 40.136      Percent Identity: 24.830
alignment_block:
  MS-00-0412-rcb

```

Align seg 1/1 to: AAB59817 from: 1 to: 999

```

658 .....GCTGGTACTCGTCGCACAAATGTTTCCAGGTTCAG 622
127 ArgTrpSerArg...CysSerSerProSerProArgArgCysProPse 144
||||| ||| ||||| ||||| |||||
624 ATGGCATTCGACGAGCACTGAGACCT.....TCGACG 586
159 rgserrargapserrelaglyproargalaalaaRpheargargcysarg 175
::: ||||| ||||| |||||
595 ACTGCCGTC.....GTCTCGGATATTCACACAGTCGTCG 561
176 AspAlaCysGslIarGArGalAaArgCysProGlyProArgSerAlaProSe 192
||| ||| ::::::|||::|
560 TGCACACTCATGCGGCTCCGAAGAGTACTTAAGCGGACACAGCTCTCC 511
192 rleatrgargglyserlaraspargserlaralasrargserargsera 209
::: ||||| ||||| ||||| |||||
510 ACAGACGACCCGCTGCCGAGATCGATTCTGAC.....CCACCTCGGGG 467
209 rggllyserProLeucySgilaThlaThlSerCysProArGalArgArg 225
|||||::|::|::|::|::|::|
466 AGGTGGCGCAGATATATGMANAGAGTCGCGCGTGAGTGAATGGACTCTTGC 417
226 ArgCysSerIle.....GlyAlaSerSergLy..... 234
|||||::|::|::|::|::|
416 ATGCACTGTACACTCCTCCA..... 395
235 .....CysProHisProProValArgArgSerProValanserSerL 249
||||| ||||| |||||
394 .....GCSTATTGTACATAGCCCTGCTCCACGAGCGGCTTGTGCGCANGT 350
249 ysAtgAlaHisAlaArgCysThrAlaAlaArgGlyIarGpheArgGlyPro 265
||||| ||| ||| ::||| ||||| |||
349 TTCTCCAGCAACCATTACGCAAGCC..... 326
266 ThrSerIarGAspTrnGlyIarGAlaArgCysTrpParGTrProArgProAr 282
::: |||||::|::|::|::|
325 .....TAAAATAGACGCGGCTCTTATTTTACTCTT 292
282 gArGcysArgCysSerArGArGTrPrGlyAlaArgProLeu..... 294

```

```

291 GTGTGGGGGGCGGTGGCGCATGCGCATGGCTGACCGCTGACATACTT 242
    ||| ||| ||| ||| ||| |||
295 ..TTPAlaSerGlyCysProAlaArgAlaArgTPrArgArgGlySerAsn... 309
    ||| ||| ||| ||| ||| |||
241 GAAGGGCATGGCCCACTCTCTGGATCTTTTACATAGCTCCACAGA 1322
    ..... |||::: |||::: |||::: |||:::
310 ..... |||::: |||::: |||::: |||:::
191 TTTTACCCCGTTGACGTTGGCCCACTTTCAGAGGGCGCTGTCA... GCAG 145
    : |||::: ||| : ::|||::: |||::: |||:::
319 r...ProLysArgThrcys.....GlyArgArgAlaArgSerAspHis 333
    : |||::: |||::: |||::: |||::: |||::: |||:::
144 GCAAATGCTGCTTCTTACTTCTTCGAAGTGGTTCAAATAGCGTCTTCC 95
    : |||::: |||::: |||::: |||::: |||::: |||:::
333 eAlaAlaArgArgSerArgCysProAlaSerSerProIleArgTPrHngly 349
    : |||::: |||::: |||::: |||::: |||::: |||:::
94 CGCTGCCGA.....TG 84
    ||| ||| ||| ||| ||| |||
350 ArgCysArgArgTPrArgArgProLeuSerGlyCysSerProAlaAlaThrcy 366
    ||| ||| ||| ||| ||| |||
83 TTTCCTCATGATGAGACGGTGAAGGCGTGGTCCCTCGCGCTACTTGGT 34
    : ||| ||| ||| ||| ||| |||
366 sTrAlaAlaArgCysGly.....ArgAspGlyCysSerAlaPhePhcGlyA 381
    : ||| ||| ||| ||| ||| |||
33 CCCCTTTCGGGCACAGATGCGTGGCTCCGCA 2
    :|||::: |||::: |||::: |||::: |||::: |||:::
381 snProLeu...HisArgSerLeuArgGlyPro 390

```

```
seq_name: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:AAW186653
seq_documentation_block:
ID: AAW18663
seq_start: 1
seq_end: 1000
seq_type: DNA
seq_format: FASTA
seq_quality: 1
seq_checksum: 1
```

standard; protein; 387 AA

AAW18663;
AC
YY

24-JUL-1997 (first entry)

Fragmented human NF-H gene +2 frameshift mutant product

KW detection; antibody; probe; cancer; neoplasia; neurodegenerative;
KW Parkinson's; Alzheimer's disease; Pick's; Huntington's disease;
KW Down's syndrome; frontal lobe dementia; progressive supranuclear palsy
KW PSP; amyotrophic lateral sclerosis; multiple sclerosis; MS;
KW cardiovascular; rheumatoid arthritis.

OS Homo sapiens.
XY

key	location/qualifiers
FT	1..387
FT	Misc-difference
FT	"/note=" "X corresponds to a stop codon in the accompanying DNA file, AA169796"
XX	

PN W09712992-A2
XX

PD 10-APR-1997
xy

PF 02-OCT-1996; 96WO-IB01106

11-JAN-1996; 96US-0009832

PR 02-001-1995; 95GB-0020080.
XX

PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
(IIYBO-) FINIV BOETENBOIV BR.

(UYUT-) UNIV STATE UTRECHT.

Burbach JPH, Grosveld FG, Van Leeuwen FW

WPI; 1997-226235/20.

N-PSDB; AAT69795.

Use of mutant genes having frame:shift mutation(s) - for developing probes for the 3'-UTR


```

466 AGCTGGCAGATATATGATAGCTCCGCTGAGTGAATGACTCTTCG 417
      |||||
819 ArgCysSerIle.....GlyAlaSerSerGly..... 827
416 ATGAACTGTACACCTCCCA..... 395
      |||||
828 .....CysProHisProProValAlaArgSerProValAsnSerSerL 842
394 .....GCCTATTGTACATGCCCTGCTCCAGCAGCCGTTGCTGCAGAT 350
      |||||
842 ysaArgAlaHisArgArgCysThrAlaArgAlaGlyArgPheArgGlyPro 858
349 TCTCCACGAGCAATAGCGAGCGC..... 326
      |||||
859 ThrSerArgAspThrGlyArgArgArgCysTrpArgTrpProArgProAr 875
325 .....TAAATGAGCGCTCCYTTATTATTGACTTCTT 292
      |||||
875 gArgCysArgCysSerArgArgTrpGlyArgProLeu..... 887
291 GTTGCTGGCGCGGTGCGACGACGACGACGACGACGACGACGACGAC 242
      |||||
888 ..TrpAlaSerGlyCysProArgAlaArgTrpArgArgGlySerAsn... 902
241 GAAGGCGCATGGCCCACTCTTGGGATCTTTGATCATCAGCTCCACGACA 192
      |||||
903 .....TrrSerSerGlyArgSerSerAlaAlaSer 912
191 TTTAACCCGTTGACGTTGCCCACTCTCGACGGGCGTCCGCTCA...GCAG 145
      |||||
912 r...ProLysArgThrCys.....GlyArgArgValArgSerAspThrS 926
144 GCAATGTCGTTGACTTCTCGAAGTGTTCATTAATACGTCGCTTCC 95
      |||||
926 eAlaArgArgSerArgCysProAlaSerSerProIleArgTrpThrGly 942
94 CGCTGCCCA.....TG 84
      |||||
943 ArgCysArgArgTrpArgArgProLeuGlyCysSerProArgAlaThrCys 959
83 TTGCCCTGATGAGGAGCGTGAAGGGCTGGGCTCCCTCGCGCATTTGTT 34
      |||||
939 sThrAlaArgCysGly.....ArgAspGlyCysSerAlaPhePheGly 974
33 CCCCTTTGGGCGACAGATGCTCCCTCCGCA 2
      |||||
974 snProLeu...HisArgSerLeuArgGlyPro 983
seq_name: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:AAW18658
seq_documentation_block:
ID AAW18658 standard; Protein: 319 AA.
AC AAW18658;
XX
XX
XX 23-JUN-1997 (first entry)
DE Fragmented human NF-L gene +2 frameshift mutant product.
XX
XX Frameshift mutation product; GAGC motif; somatic mutation; diagnosis;
XX detection; antibody; probe; cancer; neoplasia; neurodegenerative;
XX Parkinson's; Alzheimer's disease; Pick's; Huntington's disease;
XX Down's syndrome; frontal lobe dementia; progressive supranuclear palsy;
XX PSP; amyotrophic lateral sclerosis; multiple sclerosis; MS;
XX cardiovascular; rheumatoid arthritis; NF-L; neurofilament-low.
XX
XX Homo sapiens.
XX
XX OS
XX
XX Key Location/Qualifiers
XX FH Misc-difference 1..319
XX FT /note- "X corresponds to a stop codon in the
XX FT accompanying DNA file, AAT69794"

```

```

XX
XX W09712992-A2.
XX
XX 10-APR-1997.
XX
XX 02-OCT-1996; 96MO-IB01106.
XX
XX 11-JAN-1996; 96US-0009832.
XX 02-OCT-1995; 95GB-0020080.
XX
XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
XX (UYRO-) UNIV ROTTERDAM ERASMUS.
XX (UYUT-) UNIV STATE UTRECHT.
XX
XX Burbach JPH, Grosveld FG, Van Leeuwen FW.
XX
XX WPI; 1997-226235/20.
XX N-PSDB; AAT69793.
XX
XX Use of mutant genes having frame:shift mutation(s) - for developing
XX prods. for the diagnosis, prevention and treatment of associated
XX diseases, e.g. cancer or neurodegenerative disease
XX
XX Claim 22; Fig 7; 123pp: English.
XX
XX AAW18657 and AAW18658 are +1 and +2 frameshift mutations, respectively,
XX of a sequence comprising fragments of the coding sequence of the
XX human neurofilament subunit NF-L (NF-L) gene corresponding to
XX nucleotides (nt) 1-420 followed by nt 901-1440 of the wild-type NF-L
XX gene. These regions of the gene contain GAGC motifs. Frameshift mutants
XX of the tau, tubulin, apolipoprotein E, microtubule-associated protein
XX 2 (MAP-2), neurofilament subunit L, M and H and amyloid A4 genes are
XX claimed. All these genes share a common GAGC motif (N- A, G, C or T),
XX which is the site of common GA dinucleotide deletion(s) that cause
XX CC neurodegenerative disorders. Antigenic peptides used for the production
XX of antibodies, and small nucleic acid sequences derived from frameshift
XX mutants are used in the diagnosis, prevention and treatment of cancer
XX disease, Down's syndrome, frontal lobe dementia (Pick's disease),
XX CC Huntington's disease, multiple sclerosis, and other degenerative
XX diseases such as cardiovascular disease and rheumatoid arthritis.
XX
XX SQ Sequence 319 AA:

```

```

alignment_scores:
Quality: 121.50 Length: 285
Ratio: 1.038 Gaps: 14
Percent Similarity: 41.053 Percent Identity: 25.614

```

```

alignment_block:
US-09-416-579a-1 x AAW18658 ..

```

```

Align seg 1/1 to: AAW18658 from: 1 to: 319

```

```

35 CCAAGTACGCGAGGAGGACCCGCTTACGCTCTCATGAGGCGAC 84
      ||||| ||| |||||
64 ProSerLeuProCysProLeuSerProGlySerHisArgArgGlyAl 80
85 ATCGCAGCGGAGGAGACCGATATTGAACACATTCGAGAAATGACAGA 134
      |||||
80 aProAlaAlaAsnGln**ValProSerAlaThrSerArgThrThrArg. 96
135 CGCATTTGCTGCGTGCAGCGACCGCTGAGAAAGGCGGCGACGTCACG 184
      |||||
97 .....ProProThrSerGlyAlaThrThrArg 105
185 GGSTAAATCTGCTGAGCTGATGACAAAG..... 214
      ||| ||| |||||
106 ArgProGlyCysTrpSerAlaCysAlaAlaAlaThrAlaProHisAlaG 122
215 .....ATCCCAAGAGTGGCGCATGCC 236

```